Regularization by White Noise for Stochastic Functional Differential Equations

Von der Fakultät für Mathematik und Informatik
der Universität Leipzig
angenommene

Dissertation

zur Erlangung des akademischen Grades

Doctor rerum naturalium
(Dr. rer. nat.)

im Fachgebiet

Mathematik

Vorgelegt

von Dipl.-Math. Stefan Bachmann

geboren am 10.12.1989 in Leipzig

Die Annahme der Dissertation wurde empfohlen von:

1. Prof. Dr. M.-K. von Renesse (Universität Leipzig)
2. Prof. Dr. F. Flandoli (Scuola Normale Superiore di Pisa)

Die Verleihung des akademischen Grades erfolgt mit Bestehen
der Verteidigung am 06.12.2019 mit dem Gesamtprädikat magna cum laude

Contents

1. Introduction

Regularization by white noise is broadly speaking the phenomenon that certain properties of (partial) differential equations get improved by adding white noise. Usual questions concerned are existence, uniqueness and stability. It is well-known, for example, that the following ordinary differential equation

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = -\operatorname{sgn}(x(t))$$

with

$$\operatorname{sgn}(x) = \begin{cases} 1 & \text{if } x \geq 0, \\ -1 & \text{if } x < 0. \end{cases}$$

does not admit a global solution. As a matter of fact, even local existence and uniqueness are not given for bounded, measurable drifts in general. However, adding a white noise $\varepsilon \mathrm{d}W(t)/\mathrm{d}t$ to the right hand side for arbitrary small $\varepsilon > 0$ regularizes the equation in the sense that there exists a global, (pathwise) unique solution. This regularization effect has been studied vastly by a broad community and we name only a few. Portenko (cf. [25]), Veretennikov (cf. [31]) and Zvonkin (cf. [39]) have considered well-posedness for stochastic differential equations (SDEs) with bounded, discontinuous drifts. Meyer-Brandis and Proske have provided a direct approach for the construction of strong solutions in [22] based on the Malliavin calculus. For locally unbounded drifts and additive noise, Krylov and Röckner showed pathwise uniqueness and existence in [19]. In their work, the drift has to fulfill an integrability assumption based on mixed norms, which is sometimes called the Krylov-Röckner condition, see (2.2). Multiplicative noise, but drifts with stricter integrability assumptions, have been studied by Martínez and Gyōngy in [16]. Zhang extended these results to the case of drifts fulfilling the Krylov-Röckner condition and multiplicative noise in [37]. Path-by-path uniqueness for discontinuous, bounded drifts and additive noise has been proven by Davie in [7] and [8]. Different perturbing noises have been studied, for example, by Zhang (cf. [38]) or Nualart and Ouknine (cf. [23]). Catellier and Gubinelli introduced a notion of irregularity of paths in [4] to give a pathwise characterization of regularizing noises. Pilipenko and Proske have addressed the selection problem for vanishing noise in [24]. Gess and Maurelli have proven well-posedness by noise for conservation laws in [15].

We consider the regularization effect for stochastic functional differential equations and are interested in the problem of well-posedness and the strong Feller property. Well-posedness has been shown in [1] and [17], where the drift consists of a continuous, functional part and a non-functional, locally unbounded part. The difficulty for the strong Feller property lies in the state space $C\left([-r, 0], \mathbb{R}^d\right)$ of path segments with some

$r > 0$, which makes PDE methods more or less unavailable in contrast to the non-functional case where one considers the Euclidean space. Es-Sarhir, von Renesse and Scheutzow established a Harnack-inequality for additive noise in [11] whereas Wang and Yuan proved a log-Harnack inequality for multiplicative noise in [33]. Both works rely on a coupling technique which is unavailable for discontinuous drifts. To address these issues, we have developed a convergence concept for random variables in [2] and applied it to functional SDEs to derive a mostly probabilistic method for proving the strong Feller property. Additionally, it extends the probabilistic approach of Maslowski and Seidler in [21]. In [2], the strong Feller property has been shown for discontinuous, functional drifts with a sublinear growth condition. Well-posedness and the strong Feller for locally unbounded, discontinuous, functional drifts have been proven in [3]. For that integrability assumptions have been considered which are analogous to the non-functional case.

In this work, we present our main results, the basic ideas and components to prove well-posedness and the strong Feller property. All three papers [1], [2] and [3] are based on that concept and are attached in the appendix.

In chapter two, we give our main results for functional SDEs from [1] and [2]. In both papers, the drift is split up in two parts: a functional part which is assumed to growth at most sublinear and a locally unbounded, non-functional part. For proving well-posedness, we assume the functional part to be Lipschitz continuous and combine Zvonkin's transformation (cf. [39]) with a stochastic Gronwall lemma from Scheutzow and von Renesse (cf. [32]). For the non-functional part, we assume the Krylov-Röckner condition as in [19] or [37].

The main results from [3] for locally unbounded, discontinuous, functional drifts are shown in chapter three. We attempt to work only with integrability conditions analogously to the non-functional case such that it fits conceptually in the framework of regularization by white noise. The proof for well-posedness is a combination of Zvonkin's transformation to remove the non-functional part and the convergence concept to deal with the functional part.

In chapter four, we present one of our main methods, namely the convergence concept for random variables. It is mostly probabilistic and uses only a few topological arguments. Conceptually, it can be seen as the counterpart to pointwise convergence of probability measures if one interprets the convergence in probability as counterpart to the weak convergence of probability measures. Here, we only present the most important result, namely for metric spaces. In [2], a more general topological framework is considered, and a measure theoretic version is given, too. We apply it to functional SDEs, in particular to prove the strong Feller property and well-posedness. We are confident that it can also be applied to SPDEs in a similar fashion. Additionally, we study the uniform version of this convergence at the end of that chapter.

In chapter five, we discuss the basic components, ideas and methods for the main results. To avoid unnecessary technicalities and to keep it as simple as possible, we only consider toy examples, which are accessible with little previous knowledge.

Acknowledgment

I would like to express my deep gratitude to my advisor Prof. Max-Konstantin von Renesse for his great support, enthusiasm, confidence and the invaluable opportunities. I am also thankful for our stimulating discussion, his constructive critique and for sharing his deep knowledge and intuitions with me.

I also want to thank my colleagues Tobias Lehmann, Tobias Weihrauch, Dr. Vitalii Konarovskyi and Dr. Giovanni Conforti for the enjoyable atmosphere and the fruitful collaboration.

Last but not least I am sincerely grateful for my family and friends for their unlimited support, confidence and patience, the great experiences and for being an inexhaustible source of strength to me.

1.1. Notation

Notation 1.1. We denote by $\|\cdot\|_{HS}$ the Hilbert-Schmidt norm for matrices $A \in \mathbb{R}^{d\times d}$, i.e.

$$\|A\|_{HS} = \sqrt{\sum_{i,j=1}^{d} |A^{i,j}|^2}.$$

Additionally, we write for $a, b \in [-\infty, +\infty]$

$$a \wedge b := \min\{a, b\},\ \ a \vee b := \max\{a, b\}.$$

Notation 1.2. In the sequel, let $r > 0$ be an arbitrary but fixed number and define

$$\mathcal{C} := C\left([-r, 0], \mathbb{R}^d\right)$$

equipped with the supremum norm $\|\cdot\|_\infty$. For a process X defined on $[t-r, t]$ with $t \geq 0$, we write

$$X_t(s) := X(t+s),\ \ s \in [-r, 0].$$

Notation 1.3. We introduce the following function spaces: define for $0 \leq S \leq T < \infty$ and $p, q \in (1, \infty)$

$$\begin{aligned} L_p^q(S,T) &:= L^q\left([S,T]; L^p\left(\mathbb{R}^d\right)\right), & L_p^q(T) &:= L_p^q(0,T),\\ \mathbb{H}_{2,p}^q(S,T) &:= L^q\left([S,T]; W^{2,p}\left(\mathbb{R}^d\right)\right), & \mathbb{H}_{2,p}^q(T) &:= \mathbb{H}_{2,p}^q(0,T),\\ H_{2,p}^q(S,T) &:= W^{1,q}\left([S,T]; L^p\left(\mathbb{R}^d\right)\right) \cap \mathbb{H}_{2,p}^q(,T), & H_{2,p}^q(T) &:= H_{2,p}^q(0,T), \end{aligned}$$

equipped with the norm

$$\|u\|_{H_{2,p}^q(S,T)} := \|\partial_t u\|_{L_p^q(S,T)} + \|u\|_{\mathbb{H}_{2,p}^q(S,T)},\ \ u \in H_{2,p}^q(S,T).$$

Notation 1.4. If not stated otherwise, W will be a d-dimensional Brownian motion on some arbitrary but fixed probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and every strong solution shall be defined on this space.

However, weak solutions might be defined on different filtrated probability spaces. Therefore, we use the short hand notation $(X^x, \tilde{W}^x, \mathbb{Q}^x)$ where X^x is an adapted, continuous stochastic process, $\tilde{W}^x$ is an adapted Brownian motion, both with respect to some filtrated probability space $(\tilde{\Omega}^x, \tilde{\mathcal{F}}^x, \mathbb{Q}^x, (\tilde{\mathcal{F}}_t)^x_{t\geq 0})$, and $(X^x, \tilde{W}^x)$ solves the according equation with initial value x.

2. Main Results for Sublinear Functional Drifts

In this chapter we present our main results for functional SDEs

$$\begin{aligned} \mathrm{d}X^x(t) &= B(t, X^x_t)\mathrm{d}t + b(t, X^x(t))\mathrm{d}t + \sigma(t, X^x(t))\mathrm{d}W(t), \\ X^x_0 &= x \in \mathcal{C} \end{aligned} \tag{2.1}$$

where $B : \mathbb{R}_{\geq 0} \times \mathcal{C} \to \mathbb{R}^d$, $b : \mathbb{R}_{\geq 0} \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}_{\geq 0} \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ are measurable. In this chapter, we assume for pathwise uniqueness and stability that B is Lipschitz continuous. It generalizes the results of Zhang's work in [37] where pathwise uniqueness and the strong Feller property have been established for the non-functional case. There, mixed norms has been used to optimize previous results. In [1], we have shown pathwise uniqueness and stability, see Theorem 2.1 and in [2], the focus was the convergence theorem and the strong Feller property namely Theorem 4.1 and Theorem 2.2. The following condition is sometimes called the Krylov-Röckner condition.

Condition C1. (Int$b$) *One has for every $T > 0$*

$$b \in L^q_p(T)$$

where $p, q > 1$ are given with

$$\frac{d}{p} + \frac{2}{q} < 1. \tag{2.2}$$

Condition C2. (NonDeg) *Assume that for all $T > 0$ there exists some $C_\sigma = C_\sigma(T) > 0$ such that*

$$C_\sigma^{-1} I_{d\times d} \leq \sigma(t,x)\sigma(t,x)^\top \leq C_\sigma I_{d\times d} \ \forall t \in [0,T], x \in \mathbb{R}^d.$$

Condition C3. (Grad) *For the same $p, q \in (1,\infty)$ as in condition* (Int$b$)*, one has for the distributional gradient of σ*

$$\left|\nabla_x \sigma^{i,j}\right| \in L^q_{loc}\left(\mathbb{R}_{\geq 0}; L^p\left(\mathbb{R}^d\right)\right), \ i,j = 1,\dots,d.$$

Condition C4. (Lipσ) *Assume that for all $T > 0$ there exists some $C'_\sigma = C'_\sigma(T) > 0$ with*

$$\|\sigma(t,x) - \sigma(t,y)\|_{HS} \leq C'_\sigma |x-y| \ \ \forall t \in [0,T], x, y \in \mathbb{R}^d.$$

Condition C5. (SubLin) *For $t \in [0, r)$ the function $x \mapsto B(t,x)$ is continuous and for all $T > 0$ there exists some monotone increasing $g_T : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ with*

1. $|B(t,x)| \le g_T(\|x\|_\infty)\ \forall x \in \mathcal{C}, t \in [0,T]$,
2. $\lim\limits_{r\to\infty} g_T(r)/r = 0$.

Condition C6. (Lip$B$) *For all $T > 0$ there exists some $C_B = C_B(T) > 0$ such that*

$$|B(t,x) - B(t,y)| \le C_B \|x - y\|_\infty \quad \forall t \in [0,T], x, y \in \mathcal{C}.$$

The main results read as follows. The first one is Theorem 1.4 from appendix B.

Theorem 2.1 (Well-Posedness)**.** *Assume the conditions* (Int$b$)*,* (NonDeg)*,* (Grad)*,* (SubLin) *and* (Lip$B$)*. Then local pathwise uniqueness holds and there exists a global strong solution, which has almost surely α-Hölder continuous paths on every bounded interval for any $0 < \alpha < 1/2$. Additionally, for any $\gamma \ge 1$, $T > 0$ and $R > 0$, one has*

$$\mathbb{E}\,\|X_t^x - X_t^y\|_\infty^\gamma \le C\,\|x - y\|_\infty^\gamma\,,\ \ 0 \le t \le T, y \in B_R(x)$$

with a constant $C = C(\gamma, T, R, d, q, p, C_\sigma, C_B, \|b\|_{L_p^q(T)}\,, \|\nabla\sigma\|_{L_p^q(T)}\,, g)$ where X^x and X^y denote the strong solutions of equation (2.1) *with initial values x and y respectively.*

The convergence theorem 4.1 is crucial for the following result - Theorem 1.6 from Appendix C.

Theorem 2.2 (Strong Feller Property)**.** *Assume the conditions* (Int$b$)*,* (NonDeg)*,* (Lipσ) *and* (SubLin)*. Then for each initial value $x \in \mathcal{C}$, equation* (2.1) *has a global weak solution $(X^x, \tilde{W}^x, \mathbb{Q}^x)$, which is unique in distribution. Furthermore, one has the strong Feller property for all $t > r$, i.e.*

$$\lim_{y\to x} \mathbb{E}_{\mathbb{Q}^y} f(X_t^y) = \mathbb{E}_{\mathbb{Q}^x} f(X_t^x)\ \forall f \in B_b(\mathcal{C}).$$

Moreover, if condition (Lip$B$) *is fulfilled, then equation* (2.1) *has a unique strong solution and it holds*

$$\lim_{y\to x} \mathbb{E}_{\mathbb{P}}\,|f(X_t^y) - f(X_t^x)| = 0\ \forall f \in B_b(\mathcal{C}).$$

Now, we want to briefly discuss the necessity of the assumptions made above for the strong Feller property. We denote by real numbers as initial values the corresponding constant path.

The following example shows that the strong Feller property is not fulfilled in general if the diffusion coefficient depends on the past even with "perfect" properties, i.e. uniform non-degeneracy, boundedness and smoothness.

Example 2.3 (Functional Diffusion Coefficient)*.* Consider

$$\begin{aligned} \mathrm{d}X^x(t) &= (\arctan(X(s-r)) + 2)\,\mathrm{d}W(t),\\ X_0^x &= x. \end{aligned}$$

Then we have

$$[X^x](t) = \int_0^t \left(\arctan\left(X^x(s-r)\right) + 2\right)^2 \mathrm{d}s$$

where $[X]$ denotes the quadratic variation of X. Thus,

$$\frac{\partial}{\partial t}[X^x](t) = \left(\arctan\left(X^x(t-r)\right) + 2\right)^2.$$

If we know the path segment X_T^x, we will be able to reconstruct the initial value x. Therefore, $P_{X_T^x}$ and $\mathbb{P}_{X_T^y}$ are mutually singular if $x \neq y$. Thus the strong Feller property is not given.

It might seem that the continuity assumption from condition (SubLin) is an artificial technicality. However, the next example shows that this assumption is necessary in general.

Example 2.4. Consider

$$\begin{aligned}
\mathrm{d}X^x(t) &= \operatorname{sgn}\left(X^x(t-r)\right)\mathrm{d}t + \mathrm{d}W(t),\\
X_0^x &= x
\end{aligned}$$

with

$$\operatorname{sgn}(x) = \begin{cases} 1 & \text{if } x \geq 0,\\ -1 & \text{if } x < 0. \end{cases}$$

Then we have

$$\begin{aligned}
X^0(r) &= W(r) + r,\\
X^{-1/n}(r) &= W(r) - r - \frac{1}{n}.
\end{aligned}$$

Obviously, the strong Feller property does not hold.

3. Main Results for Locally Unbounded Functional Drifts

The claim of this chapter and the paper [3] is to formulate assumptions for functional drifts such that they fit into the framework of regularization by white noise. Therefore we give integrability conditions - analogously to the non-functional case - and try to get rid of continuity assumptions as far as possible. There, a couple of problems arises. First, the semigroup theory for the state space $\mathcal{C}$ is not as well understood as for the usual Euclidean space $\mathbb{R}^d$. Thus, the trick of removing the discontinuous functional drift by semigroup methods, see section 5.2, does not work in general. Second, since $\mathcal{C}$ is not locally compact, there exists no translation invariant Borel measure such that there are no canonical Lebesgue spaces on $\mathcal{C}$ in our context. Consequently, our methods rely heavily on purely probabilistic arguments, see section 5.1. In contrast to chapter 2, we do not use mixed norms.

$$\begin{aligned} \mathrm{d}X^x(t) &= B(t, X^x)\mathrm{d}t + \sigma(t, X^x(t))\mathrm{d}W(t), \\ X_0 &= x \in \mathcal{C} \end{aligned} \tag{3.1}$$

where W is a d-dimensional Brownian motion, $B : \mathbb{R}_{\geq 0} \times C(\mathbb{R}_{\geq -r}, \mathbb{R}^d) \to \mathbb{R}^d$ is non-anticipating and $\sigma : \mathbb{R}_{\geq 0} \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ is measurable.

Condition C7. (Int$B$) *For each $T > 0$ there exist a measurable $F : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ with*

$$\int_0^T \int_{\mathbb{R}^d} |F(t, x)|^{d+1} \,\mathrm{d}x\mathrm{d}t < \infty$$

and $C_1 = C_1(T), C_2 = C_2(T) \geq 0$ with

$$\int_0^t |B(s, x)|^2 \,\mathrm{d}s \leq \int_0^t |F(s, x(s))| \,\mathrm{d}s + C_1 \sup_{s\in[-r,t]} |x(s)|^2 + C_2$$

for all $t \in [0, T]$ and $x \in C\left(\mathbb{R}_{\geq -r}, \mathbb{R}^d\right)$.

Condition C8. (Split) *Assume that there is an $r_{\tilde{B}} \in (0, r)$ such that*

$$B(t, x) = \tilde{B}(t, x) + b(t, x(t))$$

with $b \in L^{2d+2}\left(\mathbb{R}_{\geq 0} \times \mathbb{R}^d; \mathbb{R}^d\right)$ and $\tilde{B} : \mathbb{R}_{\geq 0} \times C\left(\mathbb{R}_{\geq -r}, \mathbb{R}^d\right) \to \mathbb{R}^d$ measurable where, for fixed $t \geq 0$, $\tilde{B}(t, x)$ depends only on $x_{|[-r,t-r_{\tilde{B}}]}$, i.e.

$$\tilde{B}(t, x) = \tilde{B}(t, y) \text{ if } x(s) = y(s) \ \forall s \in [-r, t - r_{\tilde{B}}].$$

Condition C9. (Unif) *For $t \in [0, r)$ the function $x \mapsto B(t, x)$ is continuous. Moreover, for each $T > 0$ there exist functions $\tilde{F} \in L^{d+1}_{loc}\left([0, T] \times \mathbb{R}^d\right)$ and $G, H : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ with G monotone increasing and*

$$\lim_{R \to \infty} \frac{H(R)}{R} = \infty$$

such that

$$\int_0^t H\left(|B(s, x)|^2\right) \mathrm{d}s \leq \int_0^t \left|\tilde{F}(s, x(s))\right| \mathrm{d}s + G\left(\sup_{s \in [-r, t]} |x(s)|\right)$$

for all $t \in [0, T]$ and $x \in C\left(\mathbb{R}_{\geq -r}, \mathbb{R}^d\right)$.

Condition C10. (BdMem) *The non-anticipating function B has bounded memory, i.e. it holds*

$$B(t, x) = B(t, y) \text{ if } x(s) = y(s) \ \forall s \in [t - r, t].$$

Then we use the abuse of notation

$$B(t, x_t) = B(t, x) \ \forall x \in C\left(\mathbb{R}_{\geq -r}, \mathbb{R}^d\right)$$

and similarly for $\tilde{B}$ if (Split) *is satisfied.*

The main results read as follows which are theorems 1.4 - 1.7 from appendix D.

Theorem 3.1 (Existence)**.** *Assume* (NonDeg), (Lipσ) *and* (Int$B$). *Then for each initial value $x \in \mathcal{C}$, equation* (3.1) *has a global weak solution $(X^x, \tilde{W}^x, \mathbb{Q}^x)$, which is unique in distribution.*

Theorem 3.2 (Pathwise Uniqueness)**.** *Assume the localized versions of* (NonDeg), (Lipσ), (Int$B$) *and* (Split). *Then local pathwise uniqueness holds for equation* (3.1), *i.e. let (X^x, W) and $(\hat{X}^x, W)$ be two weak solutions of equation* (3.1) *with initial value $x \in \mathcal{C}$ on some time interval $[0, \tau]$ for some common Brownian motion W and stopping time τ. Then it follows $X^x = \hat{X}^x$ on $[0, \tau]$ almost surely.*

Theorem 3.3 (Strong Feller Property)**.** *Assume* (NonDeg), (Lipσ), (Int$B$), (Unif) *and* (BdMem). *Let $(X^x, \tilde{W}^x, \mathbb{Q}^x)$ be weak solutions with initial value $x \in \mathcal{C}$. Then one has the strong Feller property for all $t > r$, i.e.*

$$\lim_{y \to x} \mathbb{E}_{\mathbb{Q}^y} f(X_t^y) = \mathbb{E}_{\mathbb{Q}^x} f(X_t^x) \ \forall f \in B_b(\mathcal{C}).$$

Theorem 3.4 (Stability)**.** *Assume* (NonDeg), (Lipσ), (Int$B$),(Split), (Unif) *and* (BdMem). *Let X^x be the strong solutions with initial value $x \in \mathcal{C}$. Then one has*

$$\lim_{y \to x} \mathbb{E} \left\| X_t^y - X_t^x \right\|_\infty^\gamma = 0$$

for all $0 < \gamma < 2$ and for $t > r$

$$\lim_{y \to x} \mathbb{E} \left| f(X_t^y) - f(X_t^x) \right| = 0 \ \forall f \in B_b(\mathcal{C}).$$

The following remark shows that a rich class of functionals fulfill condition (Int$B$) and condition (Unif).

Remark 3.5.

1. Conditions (Int$B$) and (Unif) are closed under linear combinations.
2. Assume, one has
$$B(t,x_t) = \int_{-r}^{0} k(t,x(t+s))\mathrm{d}\mu(s)$$
for some Borel measure μ on $[-r,0]$. Then (Int$B$) is fulfilled if k is of at most linear growth in the second variable uniformly on $[0,r)$ and
$$k \in L^{2d+2}\left([0,T]\times\mathbb{R}^d\right) \ \forall T>0.$$
If $x \mapsto k(t,x)$ is additionally continuous for $t\in[0,r)$ then condition (Unif) will be satisfied. The assumption $\operatorname{supp}\mu \subset [-r,-r_{\tilde{B}}]$ for some $r_{\tilde{B}} \in (0,r)$ implies (Split).

Example 3.6. Consider the one-dimensional, deterministic, functional equation
$$\begin{aligned}\mathrm{d}x(t) &= B(t,x(t-1))\mathrm{d}t,\\ x_0 &= 0\end{aligned}$$
with
$$B(t,z) := \begin{cases} 8t^7 & \text{if } \ t\in[0,1], z\in\mathbb{R},\\ \mathbb{1}_{|z|\le 5}\,|z|^{-1/8} & \text{otherwise.}\end{cases}$$
The drift B fulfills conditions (Int$B$), (Split) and (Unif). Computing the solution x yields for $t\in[0,1]$
$$x(t)=t^8$$
Consequently, one has to solve
$$\begin{aligned}\mathrm{d}x(t) &= (t-1)^{-1}\mathrm{d}t,\\ x(1) &= 1\end{aligned}$$
for $t\in[1,2]$. Integrating both sides yields
$$x(t) = 1+\int_1^t (s-1)^{-1}\mathrm{d}s = \infty$$
for each $t\in(1,2)$. It follows that the equation has no global solution in contrast to its regularized version
$$\mathrm{d}X(t) = B(t,X(t-1))\mathrm{d}t + \mathrm{d}W(t)$$
although the conditions (Int$B$), (Split) and (Unif) are fulfilled.

4. Convergence Concept for Random Variables in Topological Spaces

4.1. Main Result

In this section, we discuss our convergence concept for random variables. On the first sight it might look rather abstract, but we have applied it directly to functional SDEs to prove the strong Feller property and well-posedness. Here, we do not present the convergence result in its full generality to avoid a sweeping topological framework. The main convergence result - Theorem 1.7 from appendix C - reads as follows

Theorem 4.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be some probability space and (E, d) be a metric space. Furthermore, let $X, X_n : \Omega \to E$, $n \in \mathbb{N}$ be measurable maps. Then the statement*

1. a) $\lim_{n\to\infty} \mathbb{P}^* \left(d\left(X, X_n\right) \geq \varepsilon\right) = 0 \ \forall \varepsilon > 0,$

   b) $\lim_{n\to\infty} \mathbb{P}_{X_n}\left(O\right) = \mathbb{P}_X\left(O\right)$ *for all open* $O \subset E$

implies

2. $\lim_{n\to\infty} \mathbb{E}\left|f(X) - f(X_n)\right| = 0 \forall f \in B_b(E)$

where $\mathbb{P}^*$ denotes the outer measure of $\mathbb{P}$. Additionally, if there exists some null set $N \subset \Omega$ such that $X(\Omega \setminus N)$ is separable, then the converse implication is also true.

Proof. See Theorem 1.7 in [2]. □

Corollary 4.2. *Let (E, d) be a complete, separable metric space. Let $\mathbb{P}$, $\mathbb{P}_n$, $n \in \mathbb{N}$ be probability measures on E equipped with its Borel σ-algebra. Assume*

$$\lim_{n\to\infty} \mathbb{P}_n(O) = \mathbb{P}(O) \ \forall \text{ open } O \subset E.$$

Then there exist a probability space $(\Omega, \mathcal{F}, \mathbb{Q})$ and random variables $X : \Omega \to E$ and $X_n : \Omega \to E$, $n \in \mathbb{N}$ such that

$$X \sim \mathbb{P}, \ X_n \sim \mathbb{P}_n, n \in \mathbb{N}$$

and

$$\lim_{n\to\infty} \mathbb{E}\left|f(X) - f(X_n)\right| = 0 \ \forall f \in B_b(E).$$

Proof. This is a direct consequence of Theorem 4.1 and Skorokhod's theorem, see [5]. □

4.2. Examples

These examples are taken directly from [2]. The first two examples illustrate the necessity of the conditions. Then two one-dimensional SD(D)Es has been considered to investigate the difference between the strong Feller property and its "improved" version.

1. Consider a sequence $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}^d$ that converges to some $x_0 \in \mathbb{R}^d$ with $x_i \neq x_0$ for all $i \in \mathbb{N}$. Then the deterministic random variables

$$X_n \equiv x_n$$

converge pointwise to x_0 but their laws δ_{x_n} do not converge pointwise to δ_{x_0}. In particular, it holds

$$\mathbb{E}\left|\mathbb{1}_{\{x_0\}}(X) - \mathbb{1}_{\{x_0\}}(X_n)\right| = 1.$$

On the other hand, let N be a standard Gaussian random variable and consider instead the sequence

$$Y_n := x_n + N.$$

Then one has

$$\lim_{n\to\infty} \mathbb{E}\left|f(Y_n) - f(Y)\right| = 0 \;\forall f \in B_b(\mathbb{R}).$$

2. Let $X \sim \mathcal{N}(0,1)$ be a standard Gaussian random variable and define

$$X_n := -X \sim \mathcal{N}(0,1).$$

It holds

$$\mathbb{P}(X \in A) = \mathbb{P}(X_n \in A) \;\forall n \in \mathbb{N}, A \in \mathcal{B}(\mathbb{R}).$$

Obviously, X_n does not converge to X in probability and it holds

$$\mathbb{E}\left|\mathbb{1}_{\mathbb{R}_{\geq 0}}(X) - \mathbb{1}_{\mathbb{R}_{\geq 0}}(X_n)\right| = 1 \;\forall n \in \mathbb{N}.$$

3. Now, assume we have a one-dimensional SDE that has a unique strong solution for each real initial value

$$\begin{aligned}
&\mathrm{d}X(t) = b(t, X(t))\mathrm{d}t + \sigma(t, X(t))\mathrm{d}W(t),\\
&X^x(0) = x \in \mathbb{R}
\end{aligned}$$

where W is a d-dimensional Brownian motion on some probability space, and $b : \mathbb{R}_{\geq 0} \times \mathbb{R} \to \mathbb{R}$ and $\sigma : \mathbb{R}_{\geq 0} \times \mathbb{R} \to \mathbb{R}^{1,d}$ are measurable. Assume furthermore that X has the Feller property, i.e.

$$\lim_{y\to x} \mathbb{E}f(X^y(t)) = \mathbb{E}f(X^x(t)) \;\forall f \in C_b(\mathbb{R}).$$

Then for every sequence $(x_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ with $x_n \to x$ and $t > 0$, one has

$$X^{x_n}(t) \to X^x(t) \text{ a.s.}$$

In particular, by Theorem 4.1, the strong Feller property is equivalent to

$$\lim_{n\to\infty} \mathbb{E}\,|f(X^{x_n}(t)) - f(X^x(t))| = 0 \;\forall f \in B_b(\mathbb{R}).$$

This can be seen as follows: by uniqueness, one has monotonicity for the solutions, i.e. for all $x \leq y$ holds

$$X^x(t) \leq X^y(t) \;\forall t \geq 0 \text{ a.s.}$$

On the other hand, for each sequence $(x_n)_{n\in\mathbb{N}}$ with $x_n \downarrow x$, the following limit exists

$$\tilde{X}(t) := \lim_{x_n \downarrow x} X^{x_n}(t) \;\forall t \geq 0 \text{ pointwise a.s.}$$

since all X^{x_n} are bounded from below by X^x. Thus, one has for all $t \geq 0$, $f \in C_b(\mathbb{R})$

$$\mathbb{E}f(X^x(t)) = \lim_{x_n \downarrow x} \mathbb{E}f(X^{x_n}(t)) = \mathbb{E}f(\tilde{X}(t)).$$

Additionally, it holds

$$\tilde{X}(t) \geq X^x(t) \;\forall t \geq 0 \text{ a.s.}$$

Thus, $\tilde{X}$ is a modification of X^x.

4. For the sake of overview, we embed real constants in $\mathcal{C}$ naturally. Now, let us consider the one-dimensional SDDE (with $r = 1$)

$$\begin{aligned} \mathrm{d}X^x(t) &= \operatorname{sgn}(X^x(t-1))\mathrm{d}W(t) \\ X_0 &= x \in \mathcal{C} \end{aligned}$$

where we use the convention

$$\operatorname{sgn} x = \begin{cases} 1 & \text{if } x \geq 0, \\ -1 & \text{if } x < 0. \end{cases}$$

This SDDE can be solved uniquely by constructing the solution recursively. By Levy's characterization, each solution X^x is distributed on $\mathbb{R}_{\geq 0}$ like a shifted Brownian motion, in particular

$$X_t^x \sim W_t + x(0) \;\forall t > 1.$$

It is not difficult to show that one has for $t > 1$

$$\lim_{y\to x} \mathbb{E}f(X_t^y) = \lim_{y\to x} \mathbb{E}f(W_t + y(0)) = \mathbb{E}f(W_t + x(0)) = \mathbb{E}f(X_t^x) \;\forall f \in B_b(\mathcal{C}),$$

see for example [11]. So, X has the strong Feller property with respect to the state space $\mathcal{C}$. On the other hand, one has for all $y \geq 0, x < 0$

$$\|X_2^y - X_2^x\|_\infty \geq |X^y(1) - X^x(1)| = |2W(1) + y - x| \text{ a.s.}$$

Therefore, convergence in probability is not given.

4.3. Uniform Convergence

Finally, we want to study the case when random variables converges uniformly with respect to the convergence above, i.e.

$$\lim_{n\to\infty} \sup_{f\in B_b(E),\|f\|_\infty\le 1} \mathbb{E}\,|f(X_n)-f(X)| = 0.$$

for random variables $X, X_n : \Omega \to E$, $n \in \mathbb{N}$.

In polish spaces it turns out that this convergence is equivalent to the probability of being different:

Theorem 4.3. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be some probability space and $(E, \mathcal{E})$ be a measurable space such that $\Delta := \{(e,e) : e \in E\} \in \mathcal{E} \otimes \mathcal{E}$. Furthermore, let $X, Y : \Omega \to E$ be measurable. Then one has*

$$\mathbb{P}(X \neq Y) \le \sup_{f\in B_b(E),\|f\|_\infty\le 1} \mathbb{E}\,|f(X)-f(Y)| \le 2\mathbb{P}(X \neq Y).$$

Proof. This theorem is a direct consequence of the discussion below. □

Lemma 4.4. *Let $n \in \mathbb{N}$ and $(p_{ij})_{i,j=1}^n \in \mathbb{R}_{\ge 0}^{n\times n}$. Then it holds*

$$\sum_{\substack{i,j=1,\\ i\neq j}}^n p_{ij} \le \max_{a\in\mathbb{R}^n,\|a\|_\infty\le 1} \sum_{i,j=1}^n |a_i - a_j|\, p_{ij} \le 2 \sum_{\substack{i,j=1,\\ i\neq j}}^n p_{ij}$$

and its maximum is attained at some point $\bar{a} \in \{-1,1\}^n$, i.e.

$$\max_{a\in\mathbb{R}^n,\|a\|_\infty\le 1} \sum_{i,j=1}^n |a_i - a_j|\, p_{ij} = \sum_{i,j=1}^n |\bar{a}_i - \bar{a}_j|\, p_{ij}.$$

Proof. The upper bound is clear since $p_{ij} \ge 0$, $i,j = 1,\dots,n$. Furthermore, the function

$$a \mapsto \sum_{i,j=1}^n |a_i - a_j|\, p_{ij},\ \ a \in \mathbb{R}^n$$

is a convex function, thus, its maximum on $\{x \in \mathbb{R}^n : \|x\|_\infty \le 1\}$ is attained at its extremal points $a \in \{-1,1\}^n$. Now, one can prove the lower bound by induction: for $n = 1$, one has for all $(p_{ij})_{i,j=1}^n \in \mathbb{R}_{\ge 0}^{n\times n}$

$$0 = \sum_{\substack{i,j=1,\\ i\neq j}}^1 p_{ij} = \max_{a\in\mathbb{R},\|a\|_\infty\le 1} \sum_{i,j=1}^n |a_i - a_j|\, p_{ij}.$$

Now, assume $n \in \mathbb{N}$ is given such that for all $(p_{ij})_{i,j=1}^n \in \mathbb{R}_{\ge 0}^{n\times n}$

$$\sum_{\substack{i,j=1,\\ i\neq j}}^n p_{ij} \le \max_{a\in\mathbb{R}^n,\|a\|_\infty\le 1} \sum_{i,j=1}^n |a_i - a_j|\, p_{ij}.$$

Let $(p_{ij})_{i,j=1}^{n+1} \in \mathbb{R}_{\geq 0}^{(n+1)\times(n+1)}$. Then on has

$$\max_{a\in\mathbb{R}^n, \|a\|_\infty \leq 1} \sum_{i,j=1}^{n} |a_i - a_j|\, p_{ij} = \sum_{i,j=1}^{n} |\bar{a}_i - \bar{a}_j|\, p_{ij}$$

for some $\bar{a} \in \{-1, 1\}^n$. Now, define

$$\alpha := \begin{cases} 1 & \text{if } \sum_{i=1}^n \mathbb{1}_{\bar{a}_i = 1}(p_{i(n+1)} + p_{(n+1)i}) \leq \sum_{i=1}^n \mathbb{1}_{\bar{a}_i=-1}(p_{i(n+1)} + p_{(n+1)i}) \\ -1 & \text{otherwise.} \end{cases}$$

It follows

$$\begin{aligned} \max_{a\in\mathbb{R}^{n+1}, \|a\|_\infty \leq 1} \sum_{i,j=1}^{n+1} |a_i - a_j|\, p_{ij} &\geq \sum_{i=1}^{n} |\alpha - \bar{a}_i| \left(p_{i(n+1)} + p_{(n+1)i}\right) + \sum_{i,j=1}^{n} |\bar{a}_i - \bar{a}_j|\, p_{ij} \\ &\geq \sum_{i=1}^{n} (p_{i(n+1)} + p_{(n+1)i}) + \sum_{\substack{i,j=1,\\ i\neq j}}^{n} p_{ij} \\ &= \sum_{\substack{i,j=1,\\ i\neq j}}^{n+1} p_{ij}. \end{aligned}$$

□

Lemma 4.5. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be some probability space and $(E, \mathcal{E})$ be a measurable space. Furthermore, let $X, Y : \Omega \to E$ be measurable. Then one has*

$$\begin{aligned} \sup_{\substack{A_1 \dot\cup \ldots \dot\cup A_n = E,\\ A_1,\ldots,A_n \in \mathcal{E}}} \sum_{i=1}^{n} \mathbb{P}(X \in A_i, Y \notin A_i) &\leq \sup_{f\in B_b(E), \|f\|_\infty \leq 1} \mathbb{E}\, |f(X) - f(Y)| \\ &= 2 \sup_{A\in\mathcal{E}} \left(\mathbb{P}(X \in A, Y \notin A) + \mathbb{P}\left(X \notin A, Y \in A\right)\right) \end{aligned}$$

Proof. Observe that

$$\begin{aligned} &\sup_{\substack{f:\Omega\to\mathbb{R} \text{ measurable},\\ \|f\|_\infty \leq 1}} \mathbb{E}\, |f(X) - f(Y)| \\ &= \sup_{n\in\mathbb{N}} \sup_{A_1 \dot\cup \ldots \dot\cup A_n = E} \sup_{a\in\mathbb{R}^n, \|a\|_\infty \leq 1} \mathbb{E} \left| \sum_{i=1}^{n} a_i \left(\mathbb{1}_{A_i}(X) - \mathbb{1}_{A_i}(Y)\right) \right| \\ &= \sup_{n\in\mathbb{N}} \sup_{A_1 \dot\cup \ldots \dot\cup A_n = E} \sup_{a\in\mathbb{R}^n, \|a\|_\infty \leq 1} \sum_{i,j=1}^{n} |a_i - a_j|\, \mathbb{P}\left(X \in A_i, Y \in A_j\right). \end{aligned}$$

Applying the lemma above closes the proof. □

Lemma 4.6. *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be some probability space and* $(E, \mathcal{E})$ *be a measurable space. Furthermore, let* $X, Y : \Omega \to E$ *be measurable. Then one has*

$$\sup_{A \in \mathcal{E} \otimes \mathcal{E}} \mathbb{P}\left((X,X) \in A, (X,Y) \notin A\right) \leq \sup_{f \in B_b(E), \|f\|_\infty \leq 1} \mathbb{E}\left|f(X) - f(Y)\right|$$

Proof. Let $\varepsilon > 0$ and $A \in \mathcal{E} \otimes \mathcal{E}$. Then one can find $n \in \mathbb{N}$, $A_1, \ldots A_n, B_1, \ldots B_n \in \mathcal{E}$ with $A_i \cap A_j = \emptyset$ for $i, j = 1, \ldots n, i \neq j$ such that

$$\mathbb{P}\left((X,X) \in A \Delta \bigcup_{i=1}^n A_i \times B_i\right) < \varepsilon$$

and

$$\mathbb{P}\left((X,Y) \in A \Delta \bigcup_{i=1}^n A_i \times B_i\right) < \varepsilon.$$

It holds

$$\begin{aligned}
&\mathbb{P}\left((X,X) \in A, (X,Y) \notin A\right) \\
\leq &\mathbb{P}\left((X,X) \in \bigcup_{i=1}^n A_i \times B_i, (X,Y) \notin \bigcup_{i=1}^n A_i \times B_i\right) + 2\varepsilon \\
\leq &\sum_{i=1}^n \mathbb{P}\left((X,X) \in A_i \times B_i, (X,Y) \notin A_i \times B_i\right) + 2\varepsilon \\
= &\sum_{i=1}^n \mathbb{P}\left(X \in A_i \cap B_i, Y \notin B_i\right) + 2\varepsilon \\
\leq &\sum_{i=1}^n \mathbb{P}\left(X \in A_i \cap B_i, Y \notin A_i \cap B_i\right) + 2\varepsilon \\
\leq &\sup_{f \in B_b(E), \|f\|_\infty \leq 1} \mathbb{E}\left|f(X) - f(Y)\right| + 2\varepsilon.
\end{aligned}$$

□

Lemma 4.7. *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be some probability space and* $(E, \mathcal{E})$ *be a measurable space. Furthermore, let* $\Delta := \{(e,e) : e \in E\}$ *and* $X, Y : \Omega \to E$ *be measurable. Then one has*

$$\begin{aligned}
\sup_{A \in \mathcal{E} \otimes \mathcal{E}, A \subset \Delta^c} \mathbb{P}((X,Y) \in A) &\leq \sup_{f \in B_b(E), \|f\|_\infty \leq 1} \mathbb{E}\left|f(X) - f(Y)\right| \\
&\leq 2 \sup_{A \in \mathcal{E} \otimes \mathcal{E}, A \subset \Delta^c} \mathbb{P}((X,Y) \in A).
\end{aligned}$$

Proof. By the Lemma above, one has

$$\begin{aligned}
\sup_{A\in\mathcal{E}\otimes\mathcal{E}, A\subset\Delta^c} \mathbb{P}((X,Y)\in A) &= \sup_{A\in\mathcal{E}\otimes\mathcal{E}, A\subset\Delta^c} \mathbb{P}((X,Y)\in A, (X,X)\notin A)\\
&\le \sup_{f\in B_b(E), \|f\|_\infty\le 1} \mathbb{E}\,|f(X)-f(Y)|\\
&=2\sup_{A\in\mathcal{E}} \left(\mathbb{P}(X\in A, Y\notin A)+\mathbb{P}\left(X\notin A, Y\in A\right)\right)\\
&=2\sup_{A\in\mathcal{E}} \mathbb{P}((X,Y)\in A\times A^c\cup A^c\times A)\\
&\le 2\sup_{A\in\mathcal{E}\otimes\mathcal{E}, A\subset\Delta^c} \mathbb{P}((X,Y)\in A).
\end{aligned}$$

□

5. Basic Methods and Application of the Convergence Concept

5.1. Probabilistic Approach for the Strong Feller Property

In this section, we want to illustrate the strategy of proving the strong Feller property. A comparable methodology is the one of Maslowski and Seidler, see Theorem 2.1 in [21]. However, our systematic approach directly extends the second part of their theorem since it can deal naturally with measurable coefficients. The task of showing the first part of their theorem may be simplified by our convergence concept in combination with Skorokhod's representation theorem [5]. Since we want to avoid a technical overload and additional preliminary work, we make use of a toy example with state space $\mathbb{R}^d$. Consider the equation

$$\begin{aligned} \mathrm{d}X^x(t) &= b(t, X^x(t))\mathrm{d}t + \sigma \mathrm{d}W(t), \\ X^x(0) &= x \in \mathbb{R}^d \end{aligned}$$

where W is some d-dimensional Brownian motion, $\sigma \in \mathbb{R}^{d\times d}$ is invertible, $b : \mathbb{R}_{\geq 0}\times\mathbb{R}^d \to \mathbb{R}^d$ is measurable and for every $T > 0$ there exists some $C_T \in \mathbb{R}$ with

$$\begin{aligned} \langle b(t,u) - b(t,v), u - v\rangle \leq C_T \left|u - v\right|^2 \ \forall u, v \in \mathbb{R}^d, t \in [0,T], \\ |b(t,u)| \leq C_T \ \forall u \in \mathbb{R}^d, t \in [0,T]. \end{aligned}$$

Additionally, we consider its drift-free equation

$$\begin{aligned} \mathrm{d}M^x(t) &= \sigma \mathrm{d}W(t), \\ M^x(0) &= x \in \mathbb{R}^d. \end{aligned}$$

Observe that both equations have a unique strong solution and the drift-free one depends continuously on the initial value in the sense that

$$\lim_{n\to\infty} M^{x_n}(t) = M^x(t) \text{ in probability}$$

for each sequence $(x_n)_{n\in\mathbb{N}} \subset \mathbb{R}^d$ converging to x and $t > 0$. In addition, M has the strong Feller property, i.e.

$$\lim_{y\to x} \mathbb{E}f(M^y(t)) = \mathbb{E}f(M^x(t)) \ \forall f \in B_b(\mathbb{R}^d).$$

Also, $\mathbb{P}_{X^x}$ has a Girsanov density with respect to $\mathbb{P}_{M^x}$, i.e.

$$\mathbb{E}f(X^x(t)) = \mathbb{E}\left[D^x(t)f(M^x(t))\right] \ \forall f \in B_b(\mathbb{R}^d)$$

with

$$D^x(t) = \exp\left(\int_0^t \sigma^{-1}b(s, M^x(s))^\top \mathrm{d}W(s) - \frac{1}{2}\int_0^t \left|\sigma^{-1}b(s, M^x(s))\right|^2 \mathrm{d}s\right).$$

At first, we show that X has the strong Feller property, too. Let $f \in B_b\left(\mathbb{R}^d\right)$, then one has

$$\begin{aligned}
&\mathbb{E}f(X^x(t)) - \mathbb{E}f(X^y(t)) \\
=&\mathbb{E}[D^x(t)M^x(t)] - \mathbb{E}[D^y(t)M^y(t)] \\
\leq&\mathbb{E}[D^x(t)(f(M^x(t)) - f(M^y(t)))] + \|f\|_\infty \mathbb{E}\left|D^x(t) - D^y(t)\right|.
\end{aligned}$$

By Theorem 4.1, one has

$$\lim_{y\to x} \mathbb{E}\left|f(M^x(t)) - f(M^y(t))\right| = 0$$

and in particular,

$$\lim_{n\to\infty} D^x(t)f(M^{x_n}(t)) = D^x(t)f(M^x(t)) \text{ in probability}$$

for each sequence $(x_n)_{n\in\mathbb{N}} \subset \mathbb{R}^d$ converging to x. By the dominated convergence theorem, it follows

$$\lim_{y\to x} \mathbb{E}[D^x(t)(f(M^y(t)) - f(M^x(t)))] = 0.$$

Consequently, it remains to show

$$\lim_{y\to x} \mathbb{E}\left|D^x(t) - D^y(t)\right| = 0.$$

Since one has $\mathbb{E}_\mathbb{P} D^z(t) = 1$ for all $z \in \mathcal{C}$, it suffices to show for each sequence $(x_n)_{n\in\mathbb{N}} \subset \mathcal{C}$ converging to x

$$\lim_{n\to\infty} D^{x_n}(t) = D^x(t) \text{ in probability.}$$

This can be seen as follows: by Fatou's lemma,

$$2 - \lim_{n\to\infty} \mathbb{E}_\mathbb{P}\left|D^{x_n}(t) - D^x(t)\right| = \lim_{n\to\infty} \mathbb{E}_\mathbb{P}\left(D^x(t) + D^{x_n}(t) - \left|D^{x_n}(t) - D^x(t)\right|\right) \geq 2$$

would hold and the L^1-convergence would be an immediate consequence. Therefore, it is sufficient to show

$$\lim_{y\to x} \mathbb{E}\int_0^t \left|b(s, M^y(s)) - b(s, M^x(s))\right|^2 \mathrm{d}s = 0,$$

by the martingale isometry. However, this is a direct consequence of Theorem 4.1. Finally, one ends up with

$$\lim_{y\to x} \mathbb{E}f(X^y(t)) = \mathbb{E}f(X^x(t)).$$

By Itō's formula and Gronwall's lemma, one can easily show

$$\lim_{n\to\infty} X^{x_n}(t) = X^x(t) \text{ in probability}$$

for each sequence $(x_n)_{n\in\mathbb{N}} \subset \mathbb{R}^d$ converging to x. Thus, it even follows

$$\lim_{y\to x} \mathbb{E}\,|f(X^y)(t) - f(X^x(t))| = 0$$

by Theorem 4.1. Now, we can summarize the strategy for showing the strong Feller property in a few steps without specifying the details. show that

Summary 5.1. 1. the drift-free version of the original equation has a unique strong solution M^x for each initial value x.

2. M has the strong Feller property and for every sequence $(x_n)_{n\in\mathbb{N}}$ with $x_n \to x$, one has
$$\lim_{n\to\infty} M_t^{x_n} = M_t^x \text{ in probability}$$
(without specifying the state space).

3. The equation with drift has for each initial value x a weak solution X^x that is unique in distribution.

4. For every initial value x, $\mathbb{P}_{X^x}$ has a density with respect to $\mathbb{P}_{M^x}$ such that for any sequence $(x_n)_{n\in\mathbb{N}}$ with $x_n \to x$, one has
$$\lim_{x_n\to x} \frac{\mathrm{d}\mathbb{P}_{X^{x_n}}}{\mathrm{d}\mathbb{P}_{M^{x_n}}}(M^{x_n}) = \frac{\mathrm{d}\mathbb{P}_{X^x}}{\mathrm{d}\mathbb{P}_{M^x}}(M^x) \text{ in probability.}$$
As illustrated by the previous example, Theorem 4.1 is the key element for verifying this step since one has to deal with discontinuous coefficients.

If, in addition, the equation with drift has for each initial value x a unique strong solution X^x such that for every sequence $(x_n)_{n\in\mathbb{N}}$ with $x_n \to x$, one has

$$\lim_{n\to\infty} X_t^{x_n} = X_t^x \text{ in probability,}$$

then one can apply Theorem 4.1 again to deduce the “improved” version of the Feller property.

5.2. Removing the Drift by PDE Methods

Now, we want to present the idea of another core method to handle discontinuous drift coefficients, especially for non-functional SDEs. The following procedure is sometimes called Zvonkin's transformation, see [36] or [39]. It is based on a non-linear transformation of the state space to remove the drift. For the sake of simplicity, we consider the non-functional SDE

$$\mathrm{d}X(t) = b(t, X(t))\mathrm{d}t + \mathrm{d}W(t) \tag{5.1}$$

where $b \in L_p^q(T)$ with p and q fulfilling (2.2). Now, the transformation is given as follows. By Theorem A.1, for every $0 < T \leq T_0$, there exists a solution

$$\tilde{u}(\cdot;T) \in \left(H_{2,p}^q(T_0)\right)^d$$

of the coordinatewise PDE system

$$\begin{aligned} \partial_t \tilde{u}(t,x;T) + L_t \tilde{u}(t,x;T) + b(t,x) &= 0, \\ \tilde{u}(T,x;T) &= 0 \end{aligned}$$

for all $t \in [0,T]$ and $x \in \mathbb{R}^d$ where

$$L_t v(t,x) := \frac{1}{2} \sum_i^d \partial_{ii}^2 v(t,x) + b(t,x) \cdot \nabla v(t,x), \; v \in H_{2,p}^q(T_0).$$

Additionally, it holds

$$\sup_{T \in [0,T_0]} \left\| \tilde{u}^i(\cdot;T) \right\|_{H_{2,p}^q(T)} < \infty, \; i = 1, \dots, d$$

and by the embedding Theorem A.2, there exists a uniform δ such that for all $0 \leq S \leq T$ with $T - S \leq \delta$

$$|\tilde{u}(t,x;T) - \tilde{u}(t,y;T)| \leq \frac{1}{2} |x - y|$$

for all $t \in [S,T]$ and $x, y \in \mathbb{R}^d$. Furthermore, the function

$$u(t,x;T) := \tilde{u}(t,x;T) + x$$

satisfies coordinatewise the equation

$$\begin{aligned} \partial_t u(t,x;T) + L_t u(t,x;T) &= 0, \\ u(T,x;T) &= x. \end{aligned}$$

Now, choose a suitable $\delta > 0$ like above and define

$$Y^x(t) := u(t, X^x(t); T), \; t \in [(T-\delta) \vee 0, T].$$

Suppose one can apply the Itō-formula, then Y^x solves the following SDE

$$\mathrm{d}Y^x(t) = Du(t, X^x(t))\mathrm{d}W(t), \; t \in [(T-\delta) \vee 0, T]$$

where Du denotes the Fréchet derivative. As one can see, the transformation removes the drift b in change of a new diffusion coefficient, which might have worse regularity than before. Additionally, it holds

$$\frac{1}{2} |X^x(t) - X^y(t)| \leq |Y^x(t) - Y^y(t)| \leq \frac{3}{2} |X^x(t) - X^y(t)|, \; t \in [(T-\delta) \vee 0, T].$$

Consequently, estimating differences of solutions with different initial values is more or less equivalent to estimating the differences of the transformed ones. Therefore, in view of well-posedness for equation (5.1), it can be extremely useful to consider the process Y instead of X. Unfortunately, this method has its limitations for functional SDEs since there is no such well understood semigroup theory for the state space $\mathcal{C}$. Nevertheless, we have still applied it in combination with results for functional SDEs such as the stochastic Gronwall lemmas from [32] and [27] or the convergence concept presented above.

5.3. Existence and Uniqueness in the Distributional Sense

In this section, we want to discuss our general procedure to show existence and uniqueness in the distributional sense, which should be mainly based on growth conditions for the coefficients instead of regularity assumptions. Therefore, we use a kind of localized Novikov condition to apply the Girsanov theorem. That technique is inspired by [20] and allows us to derive the densities

$$\frac{\mathrm{d}X^x_{|[0,T]}}{\mathrm{d}M^x_{|[0,T]}}$$

discussed in section 5.1. Finally, we obtain existence and uniqueness, and are able to carry estimates for the drift-free equation over to the original equation. For that, consider the equation

$$\begin{aligned} \mathrm{d}X^x(t) &= f(t, X^x)\mathrm{d}t + g(t, X^x)\mathrm{d}W(t) \\ X^x_0 &= x \in \mathcal{C} \end{aligned} \tag{5.2}$$

and its drift-free version

$$\begin{aligned} \mathrm{d}M^x(t) &= g(t, M^x)\mathrm{d}W(t) \\ M^x_0 &= x \end{aligned} \tag{5.3}$$

where $f : \mathbb{R}_{\geq 0} \times C\left(\mathbb{R}_{\geq 0}, \mathbb{R}^d\right) \to \mathbb{R}^d$ and $g : \mathbb{R}_{\geq 0} \times C\left(\mathbb{R}_{\geq 0}, \mathbb{R}^d\right) \to \mathbb{R}^{d\times d}$ are measurable and non-anticipating. Assume that the drift-free equation (5.3) has a global strong solution which is locally pathwise unique. Furthermore, suppose that

$$\mathbb{E}_{\mathbb{P}}\left[\exp\left(\int_0^T \left(g(t, M^x)^{-1} f(t, M^x)\right)^\top \mathrm{d}W^x(t) - \frac{1}{2}\int_0^T \left|g(t, M^x)^{-1} f(t, M^x)\right|^2 \mathrm{d}t\right)\right] = 1$$

for every $T > 0$.

Theorem 5.2. *Equation* (5.2) *has a global weak solution for every initial values* $x \in \mathcal{C}$.

Proof. The strong solution M^x is by definition $(\mathcal{F}_t)_{t\geq 0}$-adapted where $(\mathcal{F}_t)_{t\geq 0}$ is the augmented filtration generated by W. Next, we construct a probability measure on

$$\mathcal{F}_\infty := \sigma\left(\mathcal{F}_t : t \geq 0\right)$$

such that M^x is a global weak solution for equation (5.2). Define

$$a^x(t) := g(t, M^x)^{-1} f(t, M^x),\ t \geq 0.$$

By the assumption above

$$t \mapsto \exp\left(\int_0^t a^x(t)^\top \mathrm{d}W(s) - \frac{1}{2} \int_0^t |a^x(t)|^2 \,\mathrm{d}s \right)$$

is a martingale and by Girsanov's theorem,

$$\bar{W}(t) := W(t) - \int_0^t a^x(t) \mathrm{d}s,\ t \geq 0$$

is a Brownian motion on $[0, T]$ under the probability measure

$$\mathrm{d}\bar{\mathbb{P}}_T := \exp\left(\int_0^T a^x(t)^\top \mathrm{d}W(t) - \frac{1}{2} \int_0^T |a^x(t)|^2 \,\mathrm{d}t \right) \mathrm{d}\mathbb{P}$$

and $(M^x, \bar{W}, \bar{\mathbb{P}}_T)$ is a weak solution of (5.2) on $[-r, T]$ for each $T > 0$. Additionally, one has for $0 < T_1 < T_2$

$$\bar{\mathbb{P}}_{T_1}(A) = \bar{\mathbb{P}}_{T_2}(A)\ \forall A \in \mathcal{F}_{T_1},$$

so the probability measure on $\mathcal{F}_\infty$ uniquely defined by

$$\bar{\mathbb{P}}(A) := \mathbb{P}_T(A)\ \forall T > 0, A \in \mathcal{F}_T$$

is indeed well-defined and $(M^x, \bar{W}, \bar{\mathbb{P}})$ is a global weak solution. $\square$

Theorem 5.3. *Let $(X^x, \tilde{W}^x, \mathbb{Q}^x)$ be a weak solution of equation (5.2) on some time interval $[-r, T]$, $T > 0$ and assume, it holds*

$$\mathbb{P}\left(\int_0^T \left| g(t, X^x)^{-1} f(t, X^x) \right|^2 \mathrm{d}t < \infty \right) = 1. \tag{5.4}$$

Then one has

$$\begin{aligned} \mathbb{Q}^x_{X^x}(A) &= \mathbb{E}_{\mathbb{P}}\left[\mathbb{1}_A(M^x) \exp\left(\int_0^T a^x(t)^\top \mathrm{d}W(t) - \frac{1}{2} \int_0^T |a^x(t)|^2 \,\mathrm{d}t \right) \right], \\ a^x(t) &:= g(t, M^x)^{-1} f(t, M^x),\ t \in [0, T] \end{aligned}$$

for all measurable $A \subset C([-r, T], \mathbb{R}^d)$. In particular, uniqueness in the distributional sense will hold if (5.4) is true for each weak solution.

Proof. Let $(X^x, \tilde{W}^x, \mathbb{Q}^x)$ be a weak solution on some time interval $[0, T]$, $T > 0$. Define

$$\tau^n(\omega) := \inf\left\{ t \geq 0 : \int_0^t \left| g(t, \omega)^{-1} f(s, \omega) \right|^2 \mathrm{d}s \geq n \right\} \wedge T,\ \omega \in C([-r, T], \mathbb{R}^d),\ n \in \mathbb{N}.$$

Then the stopped process $X^{x,n}(t) := X^x(t \wedge \tau^n(X^x))$, $t \in [-r, T]$ fulfills the equation

$$\mathrm{d}X^{x,n}(t) = \mathbb{1}_{t \leq \tau^n(X^{x,n})} f(t, X^{x,n})\mathrm{d}t + \mathbb{1}_{t \leq \tau^n(X^{x,n})} g(t, X^{x,n})\mathrm{d}\tilde{W}^x$$

By construction, Novikov's condition is fulfilled. Consequently, Girsanov's theorem is applicable for

$$a^{x,n}(t) := g(t, X^{x,n})^{-1} f(t, X^{x,n}),\ t \in [0, T]$$

and

$$\tilde{W}^{x,n}(t) := \int_0^{t \wedge \tau^n(X^{x,n})} a^{x,n}(t)\mathrm{d}s + \tilde{W}^x(t),\ t \geq 0$$

is a Brownian motion with respect to the probability measure

$$\mathrm{d}\mathbb{Q}^{x,n} := \exp\bigg(-\int_0^{\tau^n(X^{x,n})} a^{x,n}(t)^\top \mathrm{d}\tilde{W}^x(t) - \frac{1}{2}\int_0^{\tau^n(X^{x,n})} |a^{x,n}(t)|^2\,\mathrm{d}t\bigg)\mathrm{d}\mathbb{Q}^x.$$

The process $X^{x,n}$ solves the equation

$$\begin{aligned}\mathrm{d}X^{x,n}(t) &= g(t, X^{x,n})\mathrm{d}\tilde{W}^{x,n}(t),\ t \in [0, \tau^n(X^{x,n})],\\ X_0^{x,n} &= x.\end{aligned}$$

Such a solution is locally pathwise unique , i.e.

$$X^{x,n}(t) = M^{x,n}(t),\ t \in [-r, \tau^n(X^{x,n})]$$

where $M^{x,n}$ is the unique strong solution of

$$\begin{aligned}\mathrm{d}M^{x,n}(t) &= g(t, M^{x,n})\mathrm{d}\tilde{W}^{x,n}(t),\\ M_0^{x,n} &= x.\end{aligned}$$

and it holds

$$\tau^n(X^{x,n}) = \tau^n(M^{x,n}) \text{ a.s.}$$

Moreover, $\mathbb{Q}^x$ and $\mathbb{Q}^{x,n}$ are equivalent. Thus,

$$\begin{aligned}
&\mathbb{Q}^x(X^x \in A)\\
&= \lim_{n\to\infty} \mathbb{Q}^x(\tau^n(X^x) = T, X^x \in A)\\
&= \lim_{n\to\infty} \mathbb{Q}^x(\tau^n(X^{x,n}) = T, X^{x,n} \in A)\\
&= \lim_{n\to\infty} \mathbb{E}_{\mathbb{Q}^{x,n}}\bigg[\mathbb{1}_{\tau^n(X^{x,n})=T}\mathbb{1}_A(X^{x,n})\exp\bigg(\int_0^T a^{x,n}(t)^\top \mathrm{d}\tilde{W}^{x,n}(t) - \frac{1}{2}\int_0^T |a^{x,n}(t)|^2\,\mathrm{d}t\bigg)\bigg]\\
&= \lim_{n\to\infty} \mathbb{E}_{\mathbb{Q}^{x,n}}\bigg[\mathbb{1}_{\tau^n(M^{x,n})=T}\mathbb{1}_A(M^{x,n})\exp\bigg(\int_0^T \big(g(t, M^{x,n})^{-1} f(t, M^{x,n})\big)^\top \mathrm{d}\tilde{W}^{x,n}(t)\\
&\quad - \frac{1}{2}\int_0^T \big|g(t, M^{x,n})^{-1} f(t, M^{x,n})\big|^2\,\mathrm{d}t\bigg)\bigg]\\
&= \lim_{n\to\infty} \mathbb{E}_{\mathbb{P}}\bigg[\mathbb{1}_{\tau^n(M^x)=T}\mathbb{1}_A(M^x)\exp\bigg(\int_0^T a^x(t)^\top \mathrm{d}W^x(t) - \frac{1}{2}\int_0^T |a^x(t)|^2\,\mathrm{d}t\bigg)\bigg]\\
&= \mathbb{E}_{\mathbb{P}}\bigg[1_A(M^x)\exp\bigg(\int_0^T a^x(t)^\top \mathrm{d}W^x(t) - \frac{1}{2}\int_0^T |a^x(t)|^2\,\mathrm{d}t\bigg)\bigg]
\end{aligned}$$

for all measurable $A \subset C([-r,T],\mathbb{R}^d)$. $\square$

At this point, one is able to derive estimates for the equation with drift and to follow the basic idea discussed in section 5.1. It remains to show the assumptions which were used above. These rely on the “good” properties of the drift-free equation and on a-priori estimates for the original equation. To obtain a-priori estimates, Krylov’s estimate (for semimartingales) is extremely useful and was more or less used in each paper. Here, we present it in the special case of Itō-processes.

Let X be a d-dimensional Itō-process of the form

$$\mathrm{d}X(t) = b(t)\mathrm{d}t + \sigma(t)\mathrm{d}W(t)$$

where W is a d-dimensional Brownian motion. Set

$$a^{ij}(t) := \frac{1}{2}\sigma(t)\sigma(t)^\top$$

and let τ_R be the first exit time of $X(t)$ from the ball B_R.

Lemma 5.4 (Krylov’s Estimate). *For every stopping time γ and nonnegative Borel function $f : \mathbb{R}_{\geq 0} \times \mathbb{R}^d \to \mathbb{R}_{\geq 0}$ one has*

$$\begin{aligned}
&\mathbb{E}\int_0^{\gamma\wedge\tau_R} (\det a(t))^{\frac{1}{d+1}} f(t,X(t))\mathrm{d}t \\
\leq &N(d)\left(\mathbb{B}^2 + \mathbb{A}\right)^{\frac{d}{2(d+1)}} \left(\int_0^\infty \int_{|x|\leq R} f^{d+1}(t,x)\mathrm{d}x\mathrm{d}t\right)^{\frac{1}{d+1}}
\end{aligned}$$

where

$$\mathbb{A} := \mathbb{E}\int_0^{\gamma\wedge\tau_R} \operatorname{tr} a(t)\mathrm{d}t, \quad \mathbb{B} := \mathbb{E}\int_0^{\gamma\wedge\tau_R} |b(t)|\,\mathrm{d}t$$

and $N(d)$ is a constant depending only on the dimension d.

For the concrete usage we refer to the different papers and to [16] or [37].

A. Appendix

Theorem A.1. *Assume conditions* (Int$b$), (NonDeg) *and* (Grad). *Then for any* $T > 0$ *and* $f \in L_p^q(T)$, *there exists a unique solution* $u \in H_{2,p}^q(T)$ *of the following PDE*

$$\begin{aligned} \partial_t u(t,x) + L_t u(t,x) + f(t,x) &= 0, \\ u(T,x) &= 0 \end{aligned}$$

with the bound

$$\|u\|_{H_{2,p}^q(S,T)} \leq C \|f\|_{L_p^q(S,T)}$$

for any $S \in [0,T]$ *and some constant* $C = C(T, C_\sigma, p, q, \|b\|_{L_p^q(T)}) > 0$.

Proof. See [37]. □

Theorem A.2. *Let* $p, q \in (1,\infty)$, $T > 0$ *and* $u \in H_{2,p}^q(T)$.

1. *If* $\frac{d}{p} + \frac{2}{q} < 2$, *then* u *is a bounded Hölder continuous function on* $[0,T] \times \mathbb{R}^d$ *and for any* $0 < \varepsilon$, $\delta \leq 1$ *satisfying*

$$\varepsilon + \frac{d}{p} + \frac{2}{q} < 2, \quad 2\delta + \frac{d}{p} + \frac{2}{q} < 2,$$

   there exists a constant $N = N(p,q,\varepsilon,\delta)$ *such that*

$$\begin{aligned} |u(t,x) - u(s,x)| &\leq N |t-s|^\delta \|u\|_{\mathbb{H}_{2,p}^q(T)}^{1-\frac{1}{q}-\delta} \|\partial_t u\|_{L_p^q(T)}^{\frac{1}{q}+\delta}, \\ |u(t,x)| + \frac{|u(t,x) - u(t,y)|}{|x-y|^\varepsilon} &\leq N T^{-\frac{1}{q}} \left(\|u\|_{\mathbb{H}_{2,p}^q(T)} + T \|\partial_t u\|_{L_p^q(T)} \right) \end{aligned}$$

   for all $s,t \in [0,T]$ *and* $x,y \in \mathbb{R}^d, x \neq y$.

2. *If* $\frac{d}{p} + \frac{2}{q} < 1$, *then* ∇u *is a bounded Hölder continuous function on* $[0,T] \times \mathbb{R}^d$ *and for any* $\varepsilon \in (0,1)$ *satisfying*

$$\varepsilon + \frac{d}{p} + \frac{2}{q} < 1,$$

   there exists a constant $N = N(p,q,\varepsilon)$ *such that*

$$\begin{aligned} |\nabla u(t,x) - \nabla u(s,x)| &\leq N |t-s|^\delta \|u\|_{\mathbb{H}_{2,p}^q(T)}^{1-\frac{1}{q}-\frac{\varepsilon}{2}} \|\partial_t u\|_{L_p^q(T)}^{\frac{1}{q}+\frac{\varepsilon}{2}}, \\ |\nabla u(t,x)| + \frac{|\nabla u(t,x) - \nabla u(t,y)|}{|x-y|^\varepsilon} &\leq N T^{-\frac{1}{q}} \left(\|u\|_{\mathbb{H}_{2,p}^q(T)} + T \|\partial_t u\|_{L_p^q(T)} \right) \end{aligned}$$

   for all $s,t \in [0,T]$ *and* $x,y \in \mathbb{R}^d, x \neq y$.

Proof. See [12, p. 22, 23, 36]. □

Well-Posedness and Stability for a Class of Stochastic Delay Differential Equations with Singular Drift

Stefan Bachmann

Email: bachmann@math.uni-leipzig.de

Institut für Mathematik, Universität Leipzig,
Augustusplatz 10, 04109 Leipzig, Germany

December 17, 2018

Abstract

In this paper we prove well-posedness and stability of a class of stochastic delay differential equations with singular drift. Moreover, we show local well-posedness under localized assumptions.

Keywords: Stochastic functional differential equation, Strong solution, Singular drift, Zvonkin's transformation, Krylov's estimate

MSC2010: 34K50

1. Introduction and Main Results

In this paper we prove well-posedness and stability results for stochastic delay differential equations of the form

$$\mathrm{d}X(t) = V(t, X_t)\mathrm{d}t + b(t, X(t))\mathrm{d}t + \sigma(t, X(t))\mathrm{d}W(t) \tag{1}$$

where $b : \mathbb{R}_{\geq 0} \times \mathbb{R}^d \to \mathbb{R}^d$ is a Borel-function and $\sigma : \mathbb{R}_{\geq 0} \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ is measurable, bounded and non-degenerate, $V : \mathbb{R}_{\geq 0} \times C\left([-r, 0], \mathbb{R}^d\right) \to \mathbb{R}^d$ is measurable and

$$X_t(s) := X(t + s),\ \ s \in [-r, 0].$$

This generalizes previous results for the non-delay case

$$\mathrm{d}X(t) = b(t, X(t))\mathrm{d}t + \sigma(t, X(t))\mathrm{d}W(t). \tag{2}$$

Krylov and Röckner showed in [6] that equation (2) has a unique strong solution, essentially assuming $|b| \in L_p^q := L^q\left(\mathbb{R}_{\geq 0}; L^p\left(\mathbb{R}^d\right)\right)$, $\sigma \equiv \mathrm{Id}$ with

$$\frac{d}{p} + \frac{2}{q} < 1 \tag{3}$$

which elaborates previous results, in particular from Zvonkin [15], Portenko [8] and Veretennikov [10]. Gyöngy and Martínez proved existence and uniqueness theorems for non-constant σ Lipschitz in space and $|b| \in L^{2d+2}\left(\mathbb{R}_{\geq 0} \times \mathbb{R}^d\right)$ in [5]. Different stochastic flow theorems were studied by Gubinelli, Priola, Flandoli and Fedrizzi (cf. [4], [3]). Zhang showed existence of a unique strong solution and flow theorems for $|b| \in L_p^q$ and σ with $|\nabla_x \sigma| \in L_{loc}^q\left(\mathbb{R}_{\geq 0}; L^p\left(\mathbb{R}^d\right)\right)$ in [13]. Additionally, he considered equations with Sobolev drifts and driven by α-stable processes (cf. [14]).

Our general approach is to remove the drift b by Zvonkin's transformation as in [13] and to combine it with different Girsanov techniques and a stochastic Gronwall lemma from von Renesse and Scheutzow in [11, 9].

Throughout this paper, the following notation will be used

Notation 1.1. Fix $r > 0$ and define

$$\mathcal{C} := C\left([-r, 0], \mathbb{R}^d\right)$$

equipped with the supremum norm $\|\cdot\|_\infty$. For a process X defined on $[t-r, t]$ with $t \geq 0$, we define

$$X_t(s) := X(t+s), \ s \in [-r, 0].$$

For a matrix $A \in \mathbb{R}^{d \times d}$, we denote by $\|\cdot\|_{HS}$ the Hilbert-Schmidt-norm:

$$\|A\|_{HS} := \sqrt{\sum_{i,j=1}^{d} |A^{i,j}|^2}$$

and for $s, t \in [-\infty, +\infty]$, we write

$$s \wedge t := \min(s, t),$$
$$s \vee t := \max(s, t).$$

The following conditions on b and σ are the same as in [13].

Condition C1.

$$|b| \in L^q\left(\mathbb{R}_{\geq 0}; L^p\left(\mathbb{R}^d\right)\right)$$

for $p, q > 1$ satisfying (3).

Condition C2. *The diffusion coefficient σ is uniformly continuous in $x \in \mathbb{R}^d$ locally uniformly with respect to $t \in \mathbb{R}_{\geq 0}$, and $\sigma\sigma^\top$ is bounded and uniformly elliptic, i.e. there exists a $\kappa > 0$ such that*

$$\kappa^{-1} I_{d \times d} \leq \sigma(t,x)\sigma(t,x)^\top \leq \kappa I_{d \times d} \ \forall x \in \mathbb{R}^d, t \in \mathbb{R}_{\geq 0}.$$

Condition C3. *For the same $p, q \in (1, \infty)$ as in condition* (C1)*, one has for the distributional gradient of σ*

$$\left|\nabla_x \sigma^{i,j}\right| \in L^q_{loc}\left(\mathbb{R}_{\geq 0}; L^p\left(\mathbb{R}^d\right)\right), \; i, j = 1, \ldots, d.$$

Now, we state our conditions on the functional drift V:

Condition C4. *The function $V : \mathbb{R}_{\geq 0} \times \mathcal{C} \to \mathbb{R}^d$ is assumed to be sublinear in the sense that there exists a monotone increasing function $g : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ with*

$$\limsup_{r \to \infty} \frac{g(r)}{r} = 0$$

such that

$$|V(t, x)| \leq g(\|x\|_\infty)$$

for all $t \in \mathbb{R}_{\geq 0}$, $x \in \mathcal{C}$.

Condition C5. *There exists some $K > 0$ such that*

$$|V(t, x) - V(t, y)| \leq K \|x - y\|_\infty$$

for all $t \in \mathbb{R}_{\geq 0}$, $x, y \in \mathcal{C}$.

Definition 1.2. Define for $0 \leq S \leq T < \infty$ and $p, q \in (1, \infty)$

$$\begin{array}{ll} L^q_p(S,T) := L^q\left([S,T]; L^p\left(\mathbb{R}^d\right)\right), & L^q_p(T) := L^q_p(0,T), \\ \mathbb{H}^q_{2,p}(S,T) := L^q\left([S,T]; W^{2,p}\left(\mathbb{R}^d\right)\right), & \mathbb{H}^q_{2,p}(T) := \mathbb{H}^q_{2,p}(0,T), \\ H^q_{2,p}(S,T) := W^{1,q}\left([S,T]; L^p\left(\mathbb{R}^d\right)\right) \cap \mathbb{H}^q_{2,p}(S,T), & H^q_{2,p}(T) := H^q_{2,p}(0,T), \end{array}$$

equipped with the norm

$$\|u\|_{H^q_{2,p}(S,T)} := \|\partial_t u\|_{L^q_p(S,T)} + \|u\|_{\mathbb{H}^q_{2,p}(S,T)}, \; u \in H^q_{2,p}(S,T).$$

Definition 1.3. Throughout this paper, we fix a standard d-dimensional Brownian motion $W = (W(t))_{t \geq 0}$ defined on some filtrated probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \geq 0})$ satisfying the usual conditions. Let X be a local $(\mathcal{F}_t)_{t \geq 0}$-adapted semimartingale which solves equation (1) on $[0, \tau]$, for some $(\mathcal{F}_t)_{t \geq 0}$-stopping time τ, with respect to some initial condition $X_0 = \xi$ for a $\mathcal{C}$-valued, $\mathcal{F}_0$-measurable random variable ξ. Then we write $X \in \mathcal{S}^\tau(\xi)$.

Our main result reads as follows.

Theorem 1.4. *Assume all conditions* (C1), (C2), (C3), (C4) *and* (C5). *Then (local) pathwise uniqueness holds and there exists a global strong solution, which has almost surely α-Hölder continuous paths on every bounded interval for any $0 < \alpha < 1/2$. Additionally, for any $\gamma \geq 1$, $T > 0$, $R > 0$ and two solutions $X \in \mathcal{S}^T(x)$, $\hat{X} \in \mathcal{S}^T(\hat{x})$ where $x, \hat{x} \in \mathcal{C}$ with $\max\left(\|x\|_\infty, \|\hat{x}\|_\infty\right) \leq R$, one has*

$$\mathbb{E}\left\|X_t - \hat{X}_t\right\|_\infty^\gamma \leq C \|x - \hat{x}\|_\infty^\gamma, \ 0 \leq t \leq T$$

with a constant $C = C(\gamma, T, R, d, q, p, \kappa, K, \|b\|_{L_p^q(T)}, \|\nabla\sigma\|_{L_p^q(T)}, g)$.

We can also formulate a localized version of our main result as follows.

Theorem 1.5. *For any $n \in \mathbb{N}$, let $B_n := \{x \in \mathbb{R}^d : |x| \leq n\}$ and assume*

1. *$b^i, \left|\nabla_x \sigma^{i,j}\right| \in L_{loc}^{q_n}\left(\mathbb{R}_{\geq 0}; L^{p_n}\left(B_n\right)\right)$, $i, j = 1, \ldots, d$ where p_n and q_n satisfy* (3),
2. *there is a sequence of $\kappa_n > 0$ such that*
$$\kappa_n^{-1} I_{d\times d} \leq \sigma(t,x)\sigma(t,x)^\top \leq \kappa_n I_{d\times d} \ \forall x \in B_n, t \in [0, n]$$
and σ is uniformly continuous in $x \in B_n$ uniformly with respect to $t \in [0, n]$,
3. *V is locally bounded on $\mathbb{R}_{\geq 0} \times \mathcal{C}$ and for every compact $\mathcal{K} \subset \mathcal{C}$ and $T > 0$, there exists a constant $C_{\mathcal{K},T}$ such that*
$$|V(t,x) - V(t,y)| \leq C_{\mathcal{K},T} \|x - y\|_\infty \ \ \forall t \in [0, T], x, y \in \mathcal{K}.$$

Then pathwise uniqueness holds and for every $x \in \mathcal{C}$, there exists a maximal local solution until an (predictable) explosion time ζ, i.e. X solves equation (1) *on $[0, \zeta)$ with $X_0 = x$ and*

$$\inf\{t \geq 0 : X_t \notin \mathcal{K}\} < \zeta \ \text{on} \ \{\zeta < \infty\}$$

for all compact $\mathcal{K} \subset \mathcal{C}$.

Finally, if σ has no space dependence, we can relax the Lipschitz condition on V as follows.

Theorem 1.6. *Assume all conditions from Theorem 1.4. If σ has no space dependence, i.e.*

$$\sigma(t,x) = \sigma(t), \ t \in \mathbb{R}_{\geq 0}, x \in \mathbb{R}^d,$$

then one can replace the Lipschitz-condition (C5) *with*

$$|V(t, x + \gamma_t) - V(t,x)| \leq K \|\gamma_t\|_\infty \ \ \forall t \geq 0, \gamma \in \mathcal{H}_t, x \in \mathcal{C}$$

for some $K > 0$ where

$$\mathcal{H}_t := \left\{\gamma \in H^1\left([-r,t], \mathbb{R}^d\right) : \gamma(s) = 0, -r \leq s \leq 0\right\},$$

$$H^1\left([-r,t], \mathbb{R}^d\right) := \left\{\gamma : [-r,t] \to \mathbb{R}^d \ \text{absolutely continuous with} \ \int_0^t |\dot{\gamma}(s)|^2 \,\mathrm{d}s < \infty\right\}$$

to obtain pathwise uniqueness and global existence of a solution of equation (1).

This kind of condition is strongly related to Malliavin-differentiable functions with bounded Malliavin-derivative, see for example [1].

Remark 1.7. An example for a discontinuous functional $V : \mathcal{C} \to \mathbb{R}^d$, which fulfills the condition stated in Theorem 1.6, is the following:

$$V(x) = \begin{cases} 0, & x \in H^1\left([-r,0],\mathbb{R}^d\right), \\ \int_{-r}^0 x(t)\mathrm{d}t & \text{otherwise.} \end{cases}$$

2. Existence

2.1. Krylov-Type Estimates for the Non-Delay-Case

In this subsection we consider the case $V \equiv 0$.

Theorem 2.1. *Assume condition* (C2) *and that b is bounded and measurable. Furthermore, let $X \in \mathcal{S}^\tau(\xi)$ for some $\mathcal{F}_0$-measurable, $\mathcal{C}$-valued random variable ξ and stopping time τ. Let $T_0 > 0$ and $p', q' \in (0,\infty)$ be given with*

$$\frac{d}{p'} + \frac{2}{q'} < 2,$$

there exists a constant $C(d, p', q', T_0, \kappa, \|b\|_\infty)$ such that for all $f \in L_{p'}^{q'}(T_0)$ and $0 \le S \le T \le T_0$,

$$\mathbb{E}\left[\int_{S\wedge\tau}^{T\wedge\tau} f(s,X(s))\mathrm{d}s \middle| \mathcal{F}_S\right] \le C\,\|f\|_{L_{p'}^{q'}(S,T)}.$$

Proof. See [13]. □

Lemma 2.2. *Assume conditions* (C1), (C2), (C3) *and let $X \in \mathcal{S}^T(\xi)$ for some $\mathcal{F}_0$-measurable, $\mathcal{C}$-valued random variable ξ and $T > 0$. Let $p', q' \in (0,\infty)$ be given with*

$$\frac{d}{p'} + \frac{2}{q'} < 2.$$

Then for every $R \ge 0$, there exists a constant $C_R = C_R(p, q, p', q', d, T, \kappa, \|b\|_{L_p^q(T)})$ such that

$$\mathbb{E}\exp\left(\int_0^T f(t,X(t))\mathrm{d}t\right) \le C_R$$

for all $f \in L_{p'}^{q'}(T)$ with $\|f\|_{L_{p'}^{q'}(T)} \le R$. Additionally, one has

$$\mathbb{E}\int_0^T f(t,X(t))\mathrm{d}t \le C\,\|f\|_{L_{p'}^{q'}(T)}$$

for some constant $C > 0$.

Proof. Consider the global strong solution M of the stochastic differential equation

$$\begin{aligned} \mathrm{d}M(t) &= \sigma(t, M(t))\mathrm{d}W(t), \\ M_0 &= \xi. \end{aligned}$$

Let $f \in L_{p'}^{q'}(T)$ with $\|f\|_{L_{p'}^{q'}(T)} \leq R$. Since the inequality for p' and q' is strict, one can choose $\delta > 1$ small enough such that

$$\frac{d\delta}{p'} + \frac{2\delta}{q'} < 2.$$

By Theorem 2.1, one has

$$\mathbb{E}\left[\int_S^T |f(t, M(t))|^\delta \,\mathrm{d}t \middle| \mathcal{F}_S\right] \leq C \|f\|_{L_{p'}^{q'}(S,T)}^\delta$$

for some constant $C(d, p', q', \kappa, \delta, T)$. Now, choose

$$\varepsilon := \frac{1}{2C \|f\|_{L_{p'}^{q'}(T)}^\delta \vee 1}.$$

By Lemma A.6 and Young's inequality, one obtains

$$\begin{aligned} \mathbb{E}\exp\left(\int_0^T |f(t, M(t))| \,\mathrm{d}s\right) &\leq \mathbb{E}\exp\left(\int_0^T \varepsilon |f(t, M(t))|^\delta \,\mathrm{d}s + C_{\varepsilon,\delta,T}\right) \\ &\leq 2e^{C_{\varepsilon,\delta,T}} \end{aligned}$$

with some $C_{\varepsilon,\delta,T} > 0$. Due to condition (C2) and

$$\frac{d}{p} + \frac{2}{q} < 1,$$

one can apply the same method from above for $t \mapsto \sigma^{-1}(t, M(t))b(t, M(t))$ to obtain

$$\mathbb{E}\exp\left(6\int_0^T \left|\sigma^{-1}b(t, M(t))\right|^2 \mathrm{d}t\right) < \infty.$$

In particular, the process $t \mapsto \sigma(t, M(t))^{-1}b(t, M(t))$ fulfills the Novikov condition and $(M, \tilde{W}, \mathbb{Q})$ is a weak solution of the equation

$$\begin{aligned} \mathrm{d}M(t) &= b(t, M(t))\mathrm{d}t + \sigma(t, M(t))\mathrm{d}\tilde{W}(t), \\ M_0 &= \xi \end{aligned}$$

with the probability measure

$$\mathrm{d}\mathbb{Q} := \exp\left(\int_0^T \sigma(t, M(t))^{-1}b(t, M(t)) \cdot \mathrm{d}W(t) - \frac{1}{2}\int_0^T \left|\sigma(t, M(t))^{-1}b(t, M(t))\right|^2 \mathrm{d}t\right) \mathrm{d}\mathbb{P}$$

and the Brownian motion

$$\tilde{W}(t) := W(t) - \int_0^t \sigma(s, M(s))^{-1} b(s, M(s)) \mathrm{d}s.$$

Due to Theorem 1.1 in [13], uniqueness in distribution holds for weak solutions of the SDE

$$\begin{aligned} \mathrm{d}X(t) &= b(t, X(t))\mathrm{d}t + \sigma(t, X(t))\mathrm{d}W(t), \\ X_0 &= \xi \end{aligned}$$

and the estimates from above provide

$$\begin{aligned} &\mathbb{E}_{\mathbb{P}} \exp\left(\int_0^T f(t, X(t))\mathrm{d}t\right) \\ =&\mathbb{E}_{\mathbb{Q}} \exp\left(\int_0^T f(t, M(t))\mathrm{d}t\right) \\ =&\mathbb{E}_{\mathbb{P}} \exp\Bigg(\int_0^T f(t, M(t))\mathrm{d}t + \int_0^T \sigma(t, M(t))^{-1} b(t, M(t)) \cdot \mathrm{d}W(t) \\ &\qquad\qquad - \frac{1}{2}\int_0^T \left|\sigma(t, M(t))^{-1} b(t, M(t))\right|^2 \mathrm{d}t\Bigg) \\ \leq& \left[\mathbb{E}_{\mathbb{P}} \exp\left(\int_0^T 2f(t, M(t))\mathrm{d}t\right)\right]^{\frac{1}{2}} \Bigg[\mathbb{E}_{\mathbb{P}} \exp\Bigg(2\int_0^T \sigma(t, M(t))^{-1} b(t, M(t)) \cdot \mathrm{d}W(t) \\ &\qquad\qquad - \int_0^T \left|\sigma(t, M(t))^{-1} b(t, M(t))\right|^2 \mathrm{d}t\Bigg)\Bigg]^{\frac{1}{2}} \\ \leq& \left[\mathbb{E}_{\mathbb{P}} \exp\left(\int_0^T 2f(t, M(t))\mathrm{d}t\right)\right]^{\frac{1}{2}} \cdot \left[\mathbb{E}_{\mathbb{P}} \exp\left(6\int_0^T \left|\sigma(t, M(t))^{-1} b(t, M(t))\right|^2 \mathrm{d}t\right)\right]^{\frac{1}{4}} \\ \leq& C_R \end{aligned}$$

for a constant $C_R = C_R(d, p, q, p', q', d, T, \kappa, \|b\|_{L_p^q(T)})$. By Theorem 2.1, one has

$$\int_0^T f(t, M(t))\mathrm{d}s \to 0 \text{ in probability}$$

if $\|f\|_{L_{p'}^{q'}(T)} \to 0$. Together with the exponential bound from above, it follows

$$\mathbb{E}\int_0^T f(t, X(t))\mathrm{d}t \to 0$$

if $\|f\|_{L_{p'}^{q'}(T)} \to 0$. Consequently, the linear operator $A : L_{p'}^{q'}(T) \to \mathbb{R}$ given by

$$f \mapsto \mathbb{E}\int_0^T f(t, X(t))\mathrm{d}t$$

is continuous, which provides the existence of the desired constant. $\square$

Remark 2.3. The previous lemma is a version of Theorem 2.2 in [13] but with relaxed assumptions on f. However, the proof is based on the pathwise uniqueness, which has been proven in [13].

2.2. Existence

From now on, we drop the assumption $V \equiv 0$.

Lemma 2.4. *Assume condition* (C2) *and consider a global weak solution* (M, W) *for the equation*

$$\mathrm{d}M(t) = \sigma(t, M(t))\mathrm{d}W(t).$$

Then for any $T > 0$ *and* $0 \leq \alpha < (2d\kappa T)^{-1}$*, it holds*

$$\mathbb{E}\exp\left(\alpha \sup_{0\leq t\leq T} |M(t)|^2\right) \leq \frac{4}{\sqrt{1-2\alpha d\kappa T}}\mathbb{E}\exp\left(\frac{\alpha}{1-2\alpha d\kappa T}|M(0)|^2\right) - 3.$$

Proof. Let M be a weak solution of the equation above. By conditioning on $M(0)$, it is sufficient to show the estimate for constant initial values $M(0) = \xi \in \mathbb{R}^d$. Now, define

$$N(t) := \int_0^t \sigma(t, M(t))\mathrm{d}W(t).$$

From (C2) it follows that for the quadratic variation of each coordinate $N^i,\ i = 1, \ldots, d$ of N

$$\langle N^i \rangle(T) \leq T\kappa.$$

By time-transformation and a simple computation using the reflection principle for Brownian motions, it follows for any $0 \leq \alpha < (2d\kappa T)^{-1}$

$$\begin{aligned}
&\mathbb{E}\exp\left(\alpha d \sup_{0\leq t\leq T}\left(\xi^i + N^i(t)\right)^2\right)\\
\leq&\mathbb{E}\exp\left(\alpha d \sup_{0\leq t\leq \kappa T}\left(\left|\xi^i\right| + |W(t)|\right)^2\right)\\
=&1 + \int_1^\infty \mathbb{P}\left[\exp\left(\alpha d\left(\left|\xi^i\right| + \sup_{0\leq t\leq\kappa T}|W(t)|\right)^2\right) \geq x\right]\mathrm{d}x\\
=&1 + \int_1^\infty \mathbb{P}\left(\left|\xi^i\right| + \sup_{0\leq t\leq\kappa T}|W(t)| \geq \sqrt{\frac{\ln(x)}{\alpha d}}\right)\mathrm{d}x\\
\leq&1 + 4\int_1^\infty \mathbb{P}\left(\left|\xi^i\right| + W(\kappa T) \geq \sqrt{\frac{\ln(x)}{\alpha d}}\right)\mathrm{d}x\\
\leq&4\mathbb{E}\exp\left(\alpha d\left(\left|\xi^i\right| + W(\kappa T)\right)^2\right) - 3\\
=&\frac{4}{\sqrt{1-2\alpha d\kappa T}}\exp\left(\frac{\alpha d}{1-2\alpha d\kappa T}\left|\xi^i\right|^2\right) - 3
\end{aligned}$$

where W is some one-dimensional Brownian motion. Finally, by Hölder's inequality, one obtains

$$\begin{aligned}
&\mathbb{E}\exp\left(\alpha \sup_{0\le t\le T} |M(t)|^2\right)\\
\le&\mathbb{E}\exp\left(\alpha \sum_{i=1}^{d} \sup_{0\le t\le T} \left(\xi^i + N^i(t)\right)^2\right)\\
\le&\sqrt[d]{\prod_{i=1}^{d}\mathbb{E}\exp\left(\alpha d \sup_{0\le t\le T} (\xi^i + N^i(t))^2\right)}\\
\le&\frac{4}{\sqrt{1-2\alpha d\kappa T}}\exp\left(\frac{\alpha}{1-2\alpha d\kappa T}|\xi|^2\right) - 3,
\end{aligned}$$

where the following inequality was used

$$\sqrt[d]{x_1\cdots x_d} + 3 \le \sqrt[d]{(x_1+3)\cdots(x_d+3)} \ \forall x_1,\dots,x_d \ge 0.$$

□

Corollary 2.5. *Assume conditions* (C1), (C2), (C3) *and consider a solution of* (1) *with* $V \equiv 0$. *Then one has for any* $T > 0$ *and* $0 \le \alpha < (4d\kappa T)^{-1}$ *the inequality*

$$\mathbb{E}\exp\left(\alpha \sup_{-r\le t\le T} |X(t)|^2\right) \le \frac{C}{\sqrt[4]{1-4\alpha d\kappa T}}\mathbb{E}\exp\left(\frac{\alpha}{1-4\alpha d\kappa T}\|X_0\|_\infty^2\right)$$

for a constant $C = C(d,p,q,\kappa,\|b\|_{L_p^q(T)})$.

Proof. By conditioning on X_0, it is sufficient to show the estimate for constant initial values $X_0 = \xi \in \mathcal{C}$. By condition (C2) and Lemma 2.2, the Novikov condition

$$\mathbb{E}\exp\left(\frac{1}{2}\int_0^T \left|\sigma(t,X(t))^{-1}b(t,X(t))\right|^2 \mathrm{d}t\right) < \infty$$

is fulfilled. Additionally, the bound is independent of the initial value ξ. Thus, under the probability measure

$$\begin{aligned}
&\mathrm{d}\mathbb{Q}\\
:=&\exp\left(-\int_0^T \sigma(t,X(t))^{-1}b(t,X(t))\cdot \mathrm{d}W(t) - \frac{1}{2}\int_0^T \left|\sigma(t,X(t))^{-1}b(t,X(t))\right| \mathrm{d}t\right)\mathrm{d}\mathbb{P},
\end{aligned}$$

the process

$$\tilde{W}(t) := W(t) + \int_0^t \sigma(s,X(s))^{-1}b(s,X(s))\mathrm{d}s$$

is a Brownian motion and X solves the equation

$$\mathrm{d}X(t) = \sigma(t,X(t))\mathrm{d}\tilde{W}(t)$$

on $[0,T]$.

It follows

$$
\begin{aligned}
&\mathbb{E}_{\mathbb{P}} \exp\left(\alpha \sup_{-r\le t\le T} |X(t)|^2\right)\\
=&\mathbb{E}_{\mathbb{Q}}\Bigg[\exp\left(\alpha \sup_{-r\le t\le T} |X(t)|^2\right)\cdot \exp\Bigg(\int_0^T \sigma(t,X(t))^{-1}b(t,X(t))\cdot \mathrm{d}W(t)\\
&\qquad\qquad\qquad\qquad\qquad\qquad + \frac{1}{2}\int_0^T \left|\sigma(t,X(t))^{-1}b(t,X(t))\right|^2 \mathrm{d}t\Bigg)\Bigg]\\
=&\mathbb{E}_{\mathbb{Q}}\Bigg[\exp\left(\alpha \sup_{-r\le t\le T} |X(t)|^2\right)\cdot \exp\Bigg(\int_0^T \sigma(t,X(t))^{-1}b(t,X(t))\cdot \mathrm{d}\tilde{W}(t)\\
&\qquad\qquad\qquad\qquad\qquad\qquad - \frac{1}{2}\int_0^T \left|\sigma(t,X(t))^{-1}b(t,X(t))\right|^2 \mathrm{d}t\Bigg)\Bigg]\\
\le&\left[\mathbb{E}_{\mathbb{Q}} \exp\left(2\alpha \sup_{-r\le t\le T} |X(t)|^2\right)\right]^{\frac{1}{2}}\cdot\Bigg[\mathbb{E}_{\mathbb{Q}}\exp\Bigg(2\int_0^T \sigma(t,X(t))^{-1}b(t,X(t))\cdot \mathrm{d}\tilde{W}(t)\\
&\qquad\qquad\qquad\qquad\qquad\qquad - \int_0^T \left|\sigma(t,X(t))^{-1}b(t,X(t))\right|^2 \mathrm{d}t\Bigg)\Bigg]^{\frac{1}{2}}\\
\le&\left[\mathbb{E}_{\mathbb{Q}} \exp\left(2\alpha \sup_{-r\le t\le T} |X(t)|^2\right)\right]^{\frac{1}{2}}\cdot\left[\mathbb{E}_{\mathbb{Q}}\exp\left(6\int_0^T \left|\sigma(t,X(t))^{-1}b(t,X(t))\right|^2 \mathrm{d}t\right)\right]^{\frac{1}{4}}\\
\le&\frac{C}{\sqrt[4]{1-4\alpha d\kappa T}}\exp\left(\frac{\alpha}{1-4\alpha d\kappa T}\|\xi\|_\infty^2\right)
\end{aligned}
$$

for a constant $C = C(d,p,q,\kappa,\|b\|_{L_p^q(T)})$ because of condition (C2), Lemmas 2.2 and 2.5. $\square$

Theorem 2.6. *Assume conditions* (C1), (C2), (C3) *and* (C4). *Then for any initial distribution $\mu \in \mathcal{P}(\mathcal{C})$, there exists a global weak solution (X,W) of* (1) *with $X_0 \sim \mu$.*

Proof. Due to Theorem 1.1 in [13], there exists a global strong solution X of the equation

$$
\begin{aligned}
\mathrm{d}X(t) &= b(t,X(t))\mathrm{d}t + \sigma(t,X(t))\mathrm{d}W(t),\\
X_0 &\sim \mu
\end{aligned}
$$

on the filtrated probability space $\left(\Omega,\mathcal{F},\mathbb{P},(\mathcal{F}_t)_{t\ge 0}\right)$ which is considered in this paper. Without loss of generality, one can assume that $\mathcal{F}$ is generated by the filtration $(\mathcal{F}_t)_{t\ge0}$ which is the augmented filtration generated by X_0 and W.

Firstly, assume that μ has bounded support. For $T > 0$, by Corollary 2.5,

$$\begin{aligned}&\mathrm{d}\mathbb{Q}_T\\ :=& \exp\left(\int_0^T \sigma(t, X(t))^{-1}V(t, X_t)\cdot \mathrm{d}W(t) - \frac{1}{2}\int_0^T \left|\sigma(t, X(t))^{-1}V(t, X_t)\right|^2 \mathrm{d}t\right)\mathrm{d}\mathbb{P}\end{aligned}$$

is a probability measure. Under $\mathbb{Q}_T$, the process

$$\tilde{W} := W(t) - \int_0^t \sigma(s, X(s))^{-1}V(s, X_s)\mathrm{d}s,\ t \geq 0$$

is a Brownian motion on $[0, T]$ and X is a weak solution of

$$\mathrm{d}X(t) = V(t, X_t)\mathrm{d}t + b(t, X(t))\mathrm{d}t + \sigma(t, X(t))\mathrm{d}\tilde{W}(t)$$

on $[0, T]$. Additionally, it holds for any $0 < T_1 < T_2$

$$\mathbb{Q}_{T_1}(A) = \mathbb{Q}_{T_2}(A)\ \forall A \in \mathcal{F}_{T_1}.$$

Now, let $\mathbb{Q}$ be the probability measure uniquely defined by

$$\mathbb{Q}(A) := \mathbb{Q}_T(A),\ T > 0, A \in \mathcal{F}_T.$$

Then $(X, \tilde{W}, \mathbb{Q})$ is a global weak solution. In the general case, choose a sequence of bounded, disjoint, measurable subsets $(A_n)_{n\in I} \subset \mathcal{C}$ with

$$\mu(A_n) > 0\ \forall n \in I$$

where I is some countable index set. Now, define the probability measures

$$\mu^n(\cdot) := \mu\left(\cdot\middle|A_n\right).$$

By the discussion from above, there exists for each $n \in I$ a probability measure $\mathbb{P}^n$ and a Brownian motion W^n such that $(X, W^n, \mathbb{P}^n)$ is a global weak solution with initial distribution μ^n. Now, define the probability measure

$$\hat{\mathbb{P}}(A) := \sum_{n\in I}\mathbb{P}^n(A)\mu(A_n),\ A \in \mathcal{F}$$

and the process

$$\hat{W}(\omega, t) := W^n(\omega, t) \text{ if } X_0(\omega) \in A_n.$$

Now, let $f : \Omega \to \mathbb{R}$ be measurable with finite moment with respect to $\hat{\mathbb{P}}$, $s \in \mathbb{R}_{\geq 0}$ and $A \in \mathcal{F}_s$. Then one has

$$\begin{aligned}&\mathbb{E}_{\hat{\mathbb{P}}}(\mathbb{1}_A f)\\ =&\sum_{n\in I}\mu(A_n)\mathbb{E}_{\mathbb{P}^n}(\mathbb{1}_A f)\\ =&\sum_{n\in I}\mu(A_n)\mathbb{E}_{\mathbb{P}^n}\left(\mathbb{1}_A\mathbb{E}_{\mathbb{P}^n}\left(f|\mathcal{F}_s\right)\right)\\ =&\sum_{n\in I}\mu(A_n)\mathbb{E}_{\mathbb{P}^n}\left(\mathbb{1}_{X_0\in A_n}\mathbb{1}_A\mathbb{E}_{\mathbb{P}^n}\left(f|\mathcal{F}_s\right)\right)\\ =&\sum_{n\in I}\mathbb{E}_{\hat{\mathbb{P}}}\left(\mathbb{1}_{X_0\in A_n}\mathbb{1}_A\mathbb{E}_{\mathbb{P}^n}\left(f|\mathcal{F}_s\right)\right).\end{aligned}$$

It follows

$$\mathbb{E}_{\hat{\mathbb{P}}}\left(f|\mathcal{F}_s\right) = \sum_{n\in I} \mathbb{1}_{X_0\in A_n}\mathbb{E}_{\mathbb{P}^n}\left(f|\mathcal{F}_s\right).$$

Since each W^n is a Brownian motion and a martingale under $\mathbb{P}^n$, one has for $0 \le s \le t$

$$\begin{aligned}
\mathbb{E}_{\hat{\mathbb{P}}}\left(\hat{W}(t)\middle|\mathcal{F}_s\right) &= \sum_{n\in I} \mathbb{1}_{X_0\in A_n}\mathbb{E}_{\mathbb{P}^n}\left(\hat{W}(t)\middle|\mathcal{F}_s\right) \\
&= \sum_{n,m\in I} \mathbb{1}_{X_0\in A_n}\mathbb{E}_{\mathbb{P}^n}\left(\mathbb{1}_{X_0\in A_m}W^m(t)\big|\mathcal{F}_s\right) \\
&= \sum_{n,m\in I} \mathbb{1}_{X_0\in A_n}\mathbb{1}_{X_0\in A_m}\mathbb{E}_{\mathbb{P}^n}\left(W^m(t)\big|\mathcal{F}_s\right) \\
&= \sum_{n\in I} \mathbb{1}_{X_0\in A_n}\mathbb{E}_{\mathbb{P}^n}\left(W^n(t)\big|\mathcal{F}_s\right) \\
&= \sum_{n\in I} \mathbb{1}_{X_0\in A_n}W^n(s) \\
&= \hat{W}(s)
\end{aligned}$$

and analogously

$$\mathbb{E}\left(\hat{W}^i(t)\hat{W}^j(t)\middle|\mathcal{F}_s\right) = \delta_{ij}(t-s)$$

where $i, j = 1, \ldots d$. Additionally, the process $\hat{W}$ is almost surely continuous and by Levy's characterization, $\hat{W}$ is a Brownian motion on $\left(\Omega, \mathcal{F}, \hat{\mathbb{P}}, (\mathcal{F}_t)_{t\ge 0}\right)$. Hence, $(X, \hat{W}, \hat{\mathbb{P}})$ is a weak solution with initial distribution $X_0 \sim \mu$. □

2.3. Exponential- and Krylov-Type Estimates for the General Case

In this section we show similar estimates like above for solutions with delay drift.

Lemma 2.7. *Assume conditions* (C1), (C2), (C3), (C4) *and consider a local solution $X \in \mathcal{S}^\tau(\xi)$ for some $\mathcal{F}_0$-measurable, $\mathcal{C}$-valued random variable ξ and a stopping time τ. Then one has for any $T > 0$ and $0 \le \alpha < (8d\kappa T)^{-1}$ the inequality*

$$\left[\mathbb{E}\exp\left(\alpha \sup_{-r\le t\le T\wedge\tau} |X(t)|^2\right)\right]^2 \le \frac{C}{\sqrt[4]{1-8\alpha d\kappa T}}\mathbb{E}\exp\left(\frac{2\alpha}{1-8\alpha d\kappa T}\|\xi\|_\infty^2\right)$$

for some constant $C = C(d, p, q, T, \kappa, \|b\|_{L_p^q(T)}, g, \xi)$.

Proof. Introduce the stopping times

$$\tau_n := \inf\{t \ge 0 : |V(t, X_t)| \ge n\} \wedge \tau \wedge T.$$

By the monotone convergence theorem, it suffices to show the inequality for the stopped processes X^{τ_n} with a uniform bound. The following technique is similar to the one used in [7, p. 286-297]. For every $n \in \mathbb{N}$, there exists a strong solution of the SDE

$$\begin{aligned} \mathrm{d}Y^n(t) &= b\left(t, Y^n(t)\right)\mathrm{d}t + \sigma\left(t, Y^n(t)\right)\mathrm{d}W(t),\ t \geq \tau_n, \\ Y(\tau_n) &= X(\tau_n). \end{aligned}$$

Now, define processes X^n by

$$X^n(t) := \begin{cases} X(t), & t \leq \tau_n, \\ Y^n(t), & t > \tau_n. \end{cases}$$

Under the probability measure

$$\begin{aligned} &\mathrm{d}\mathbb{Q}^n \\ &:= \exp\left(-\int_0^{\tau_n} \sigma(t, X^n(t))^{-1} V(t, X_t^n) \cdot \mathrm{d}W(t) - \frac{1}{2}\int_0^{\tau_n} \left|\sigma(t, X^n(t))^{-1} V(t, X_t^n)\right|^2 \mathrm{d}t\right) \mathrm{d}\mathbb{P}, \end{aligned}$$

the process

$$W^n(t) := W(t) + \int_0^{t \wedge \tau_n} \sigma(s, X^n(s))^{-1} V(s, X_s^n)\mathrm{d}s$$

is a Brownian motion and X^n is the unique strong solution of the equation

$$\begin{aligned} \mathrm{d}X^n(t) &= b(t, X^n(t))\mathrm{d}t + \sigma(t, X^n(t))\mathrm{d}W^n(t), \\ X_0^n &= X_0. \end{aligned}$$

Accordingly,

$$
\begin{aligned}
&\left[\mathbb{E}_{\mathbb{P}} \exp\left(\alpha \sup_{-r\le t\le T} |X^n(t)|^2\right)\right]^2 \\
=&\Bigg[\mathbb{E}_{\mathbb{Q}^n}\Bigg(\exp\left\{\alpha \sup_{-r\le t\le T} |X^n(t)|^2\right\} \cdot \exp\bigg\{ \int_0^{\tau_n} \sigma(t,X^n(t))^{-1}V(t,X^n_t)\cdot \mathrm{d}W(t) \\
&\qquad\qquad + \frac{1}{2}\int_0^{\tau_n} \left|\sigma(t,X^n(t))^{-1}V(t,X^n_t)\right|^2 \mathrm{d}t\bigg\}\Bigg)\Bigg]^2 \\
=&\Bigg[\mathbb{E}_{\mathbb{Q}^n}\Bigg(\exp\left\{\alpha \sup_{-r\le t\le T} |X^n(t)|^2\right\} \cdot \exp\bigg\{ \int_0^{\tau_n} \sigma(t,X^n(t))^{-1}V(t,X^n_t)\cdot \mathrm{d}W^n(t) \\
&\qquad\qquad - \frac{1}{2}\int_0^{\tau_n} \left|\sigma(t,X^n(t))^{-1}V(t,X^n_t)\right|^2 \mathrm{d}t\bigg\}\Bigg)\Bigg]^2 \\
\le&\mathbb{E}_{\mathbb{Q}^n} \exp\left(2\alpha \sup_{-r\le t\le T} |X^n(t)|^2\right)\cdot \mathbb{E}_{\mathbb{Q}^n}\exp\bigg(2\int_0^{\tau_n} \sigma(t,X^n(t))^{-1}V(t,X^n_t)\cdot \mathrm{d}W^n(t) \\
&\qquad\qquad - \int_0^{\tau_n} \left|\sigma(t,X^n(t))^{-1}V(t,X^n_t)\right|^2 \mathrm{d}t\bigg) \\
\le&\mathbb{E}_{\mathbb{Q}^n} \exp\left(2\alpha \sup_{-r\le t\le T} |X^n(t)|^2\right)\cdot \left[\mathbb{E}_{\mathbb{Q}^n}\exp\left(6\int_0^{T} \left|\sigma(t,X^n(t))^{-1}V(t,X^n_t)\right|^2 \mathrm{d}t\right)\right]^{\frac{1}{2}} \\
\le&\frac{C}{\sqrt[4]{1-8\alpha d\kappa T}}\mathbb{E}\exp\left(\frac{2\alpha}{1-8\alpha d\kappa T}\|\xi\|_\infty^2\right)
\end{aligned}
$$

for a constant $C = C(d,p,q,T,\kappa,\|b\|_{L^q_p(T)},g,\xi)$ due to condition (C4) and Corollary 2.5. □

Lemma 2.8. *Assume conditions* (C1), (C2), (C3), (C4) *and let* $X \in \mathcal{S}^\tau(\xi)$ *for some* $\mathcal{F}_0$*-measurable,* $\mathcal{C}$*-valued random variable* ξ *and a stopping time* τ *such that*

$$\mathbb{E}\exp\left(\varepsilon\|\xi\|_\infty^2\right) < \infty$$

for some $\varepsilon > 0$*. Let* $T > 0$ *and* $p', q' \in (0,\infty)$ *be given with*

$$\frac{d}{p'} + \frac{2}{q'} < 2.$$

Then there exists for every $R \ge 0$ *a constant* $C_R = C_R(\xi,p,q,p',q',d,T,\kappa,\|b\|_{L^q_p(T)},g)$ *such that*

$$\mathbb{E}\exp\left(\int_0^{T\wedge\tau} f(s,X(s))\mathrm{d}s\right) \le C_R$$

for all $f \in L^{q'}_{p'}(T)$ *and* $\|f\|_{L^{q'}_{p'}(T)} \le R$*. Additionally, one has*

$$\mathbb{E}\int_0^{T\wedge\tau} f(s,X(s))\mathrm{d}s \le C\|f\|_{L^{q'}_{p'}(T)}$$

with a constant $C > 0$.

Proof. By condition (C2), Lemma 2.7 and the assumption for the initial distribution, the Novikov condition

$$\mathbb{E}\exp\left(\frac{1}{2}\int_0^{T\wedge\tau}\left|\sigma(t,X(t))^{-1}V(t,X_t)\right|^2\mathrm{d}t\right)<\infty$$

is fulfilled. Therefore, under the probability measure

$$\begin{aligned}&\mathrm{d}\mathbb{Q}\\&:=\exp\left(-\int_0^{T\wedge\tau}\sigma(t,X(t))^{-1}V(t,X_t)\cdot\mathrm{d}W(t)-\frac{1}{2}\int_0^{T\wedge\tau}\left|\sigma(t,X(t))^{-1}V(t,X_t)\right|^2\mathrm{d}t\right)\mathrm{d}\mathbb{P},\end{aligned}$$

the process

$$\tilde{W}(t):=W(t)+\int_0^{t\wedge\tau}\sigma(s,X^n(s))^{-1}V(s,X^n_s)\mathrm{d}s$$

is a Brownian motion and X solves the equation

$$\mathrm{d}X(t)=b(t,X(t))\mathrm{d}t+\sigma(t,X(t))\mathrm{d}\tilde{W}(t),$$

on $[0,T\wedge\tau]$.

It follows

$$\begin{aligned}&\mathbb{E}_{\mathbb{P}}\exp\left(\int_0^{T\wedge\tau}f(t,X(t))\mathrm{d}t\right)\\
=&\mathbb{E}_{\mathbb{Q}}\Bigg[\exp\left(\int_0^{T\wedge\tau}f(t,X(t))\mathrm{d}t\right)\cdot\exp\Bigg(\int_0^{T\wedge\tau}\sigma(t,X(t))^{-1}V(t,X_t)\cdot\mathrm{d}W(t)\\
&\qquad+\frac{1}{2}\int_0^{T\wedge\tau}\left|\sigma(t,X(t))^{-1}V(t,X_t)\right|^2\mathrm{d}t\Bigg)\Bigg]\\
=&\mathbb{E}_{\mathbb{Q}}\Bigg[\exp\left(\int_0^{T\wedge\tau}f(t,X(t))\mathrm{d}t\right)\cdot\exp\Bigg(\int_0^{T\wedge\tau}\sigma(t,X(t))^{-1}V(t,X_t)\cdot\mathrm{d}\tilde{W}(t)\\
&\qquad-\frac{1}{2}\int_0^{T\wedge\tau}\left|\sigma(t,X(t))^{-1}V(t,X_t)\right|^2\mathrm{d}t\Bigg)\Bigg]\\
\leq&\left[\mathbb{E}_{\mathbb{Q}}\exp\left(2\int_0^{T\wedge\tau}f(t,X(t))\mathrm{d}t\right)\right]^{\frac{1}{2}}\cdot\Bigg[\mathbb{E}_{\mathbb{Q}}\exp\Bigg(2\int_0^{T\wedge\tau}\sigma(t,X(t))^{-1}V(t,X_t)\cdot\mathrm{d}\tilde{W}(t)\\
&\qquad-\int_0^{T\wedge\tau}\left|\sigma(t,X(t))^{-1}V(t,X_t)\right|^2\mathrm{d}t\Bigg)\Bigg]^{\frac{1}{2}}\\
\leq&\left[\mathbb{E}_{\mathbb{Q}}\exp\left(2\int_0^{T\wedge\tau}f(t,X(t))\mathrm{d}t\right)\right]^{\frac{1}{2}}\cdot\left[\mathbb{E}_{\mathbb{Q}}\exp\left(6\int_0^{T\wedge\tau}\left|\sigma(t,X(t))^{-1}V(t,X_t)\right|^2\mathrm{d}t\right)\right]^{\frac{1}{4}}\\
\leq&C_R\end{aligned}$$

where $C_R = C_R(\xi, p, q, p', q', d, T, \kappa, \|b\|_{L^q_p(T)}, g) \geq 0$ because of Lemma 2.2, condition (C4) and Lemma 2.7. For the last statement, one can use

$$\int_0^{T\wedge\tau} f(s, X(s))\mathrm{d}s \to 0 \text{ in probability with respect to } \mathbb{Q}$$

if $\|f\|_{L^{q'}_{p'}(T)} \to 0$, and the bound from above, to conclude

$$\mathbb{E}_{\mathbb{P}} \int_0^{T\wedge\tau} f(s, X(s))\mathrm{d}s \to 0$$

if $\|f\|_{L^{q'}_{p'}(T)} \to 0$. Consequently, the linear operator $A : L^{q'}_{p'}(T) \to \mathbb{R}$ given by

$$f \mapsto \mathbb{E} \int_0^{T\wedge\tau} f(s, X(s))\mathrm{d}s$$

is continuous, which provides the existence of the desired constant. □

Lemma 2.9. *Assume conditions* (C1), (C2), (C3), (C4) *and let* $X \in \mathcal{S}^\tau(\xi)$ *be a weak solution for some* $\mathcal{F}_0$*-measurable,* $\mathcal{C}$*-valued random variable* ξ *and a stopping time* τ*. Then* X *has almost surely* α*-Hölder continuous paths on* $[0, T\wedge\tau]$ *for any* $0 < \alpha < 1/2$ *and* $T > 0$*.*

Proof. Let $0 < \alpha < 1/2$ and $T > 0$.

1. $t \mapsto \int_0^{t\wedge\tau} V(s, X_s)\mathrm{d}s$ has Lipschitz continuous paths on $[0, T]$ since V is locally bounded.
2. $t \mapsto \int_0^{t\wedge\tau} b(t, X(s))\mathrm{d}s$ has almost surely α-Hölder continuous paths on $[0, T]$ because $\mathbb{E}\int_0^{T\wedge\tau} |b(t, X(t))|^2\, \mathrm{d}t < \infty$ by Lemma 2.8.
3. $t \mapsto \int_0^{t\wedge\tau} \sigma(s, X(s))\mathrm{d}W(s)$ has almost surely α-Hölder continuous paths on $[0, T]$ since σ is bounded.

□

3. Pathwise Uniqueness

Theorem 3.1. *Assume conditions* (C1), (C2), (C3), (C4), (C5)*, let* τ *be a stopping time and* $R > 0$*. For every two local solutions* $X \in \mathcal{S}^\tau(x)$ *and* $\hat{X} \in \mathcal{S}^\tau(\hat{x})$ *where* $x, \hat{x} \in \mathcal{C}$ *with* $\max(\|x\|_\infty, \|\hat{x}\|_\infty) \leq R$*, every* $\gamma \geq 1$ *and* $T_0 > 0$*, one has*

$$\mathbb{E}\left\|X_{t\wedge\tau} - \hat{X}_{t\wedge\tau}\right\|_\infty^\gamma \leq C\,\|x - \hat{x}\|_\infty^\gamma,\ \ 0 \leq t \leq T_0$$

for some constant C *depending only on* $\gamma, d, p, q, T_0, \kappa, K, \|b\|_{L^q_p(T_0)}, \|\nabla\sigma\|_{L^q_p(T_0)}, g$ *and* R*.*

By Theorem A.1, for every $0 < T \leq T_0$, there exists a solution

$$\tilde{u}(\cdot;T) \in \left(H^q_{2,p}(T_0)\right)^d$$

of the coordinatewise PDE system

$$\begin{aligned} \partial_t \tilde{u}(t,x;T) + L_t \tilde{u}(t,x;T) + b(t,x) &= 0, \\ \tilde{u}(T,x;T) &= 0 \end{aligned}$$

for all $t \in [0,T]$ and $x \in \mathbb{R}^d$ where

$$L_t v(t,x) := \frac{1}{2} \sum_{i,j,k=1}^{d} \sigma^{i,k}(t,x)\sigma^{j,k}(t,x)\partial_i \partial_j v(t,x) + b(t,x) \cdot \nabla v(t,x), \;\; v \in H^q_{2,p}(T_0).$$

Additionally, it holds

$$\sup_{T \in [0,T_0]} \left\| \tilde{u}^i(\cdot;T) \right\|_{H^q_{2,p}(T)} < \infty, \;\; i = 1, \dots, d$$

and by the embedding Theorem A.2, there exists a uniform δ such that for all $0 \leq S \leq T$ with $T - S \leq \delta$

$$|\tilde{u}(t,x;T) - \tilde{u}(t,y;T)| \leq \frac{1}{2} |x - y|$$

for all $t \in [S,T]$ and $x, y \in \mathbb{R}^d$. Furthermore, the function

$$u(t,x;T) := \tilde{u}(t,x;T) + x$$

satisfies coordinatewise the equation

$$\begin{aligned} \partial_t u(t,x;T) + L_t u(t,x;T) &= 0, \\ u(T,x;T) &= x. \end{aligned}$$

Proof. Let $X \in \mathcal{S}^\tau(x)$ and $\hat{X} \in \mathcal{S}^\tau(\hat{x})$ for some $(\mathcal{F}_t)_{t \geq 0}$-stopping time τ where $x, \hat{x} \in \mathcal{C}$. Choose $T_0 > 0$, $\gamma \geq 1$ arbitrarily and $\delta > 0$ like above. By induction, it suffices to prove for every $0 \leq S \leq T \leq T_0$ with $T - S \leq \delta$ the implication

$$\begin{aligned} \mathbb{E} \left\| X_{S \wedge \tau} - \hat{X}_{S \wedge \tau} \right\|_\infty^\gamma &\leq C_1 \left\| x - \hat{x} \right\|_\infty^\gamma \\ \Longrightarrow \mathbb{E} \left\| X_{T \wedge \tau} - \hat{X}_{T \wedge \tau} \right\|_\infty^\gamma &\leq C_2 \left\| x - \hat{x} \right\|_\infty^\gamma \end{aligned}$$

for constants C_1 and C_2 depending only on $\gamma, d, p, q, \kappa, K, T_0, \|b\|_{L^q_p(T_0)}, \|\nabla \sigma\|_{L^q_p(T_0)}, g$ and R. For the sake of simplicity, we write $u(\cdot) := u(\cdot;T)$. Furthermore, define

$$\begin{aligned} Y(t) &:= u(t, X(t)), \;\; S \wedge \tau \leq t \leq T \wedge \tau \\ \hat{Y}(t) &:= u(t, \hat{X}(t)), \;\; S \wedge \tau \leq t \leq T \wedge \tau. \end{aligned}$$

By the choice of δ, one has for the difference processes $Z(t) := X(t) - \hat{X}(t)$ and $\tilde{Z}(t) := Y(t) - \hat{Y}(t)$

$$\frac{1}{2}\left|\tilde{Z}(t)\right| \leq |Z(t)| \leq \frac{3}{2}\left|\tilde{Z}(t)\right|, \ S \wedge \tau \leq t \leq T \wedge \tau.$$

Due to Lemma 2.8, Lemma A.3 is applicable, which gives

$$\begin{aligned}\tilde{Z}(t) = &\int_{S\wedge\tau}^{t\wedge\tau} \Big(Du(s, X(s))V(s, X_s) - Du(s, \hat{X}(s))V(s, \hat{X}_s)\Big)\,\mathrm{d}s \\ &+ \int_{S\wedge\tau}^{t\wedge\tau} \Big(Du(s, X(s))\sigma(s, X(s)) - Du(s, \hat{X}(s))\sigma(s, \hat{X}(s))\Big)\,\mathrm{d}W(s)\end{aligned}$$

and consequently

$$\begin{aligned}&\mathrm{d}\left|\tilde{Z}\right|^{2\gamma}(t) \\ =&2\gamma\left|\tilde{Z}(t)\right|^{2\gamma-2}\tilde{Z}(t)^\top\Big(Du(t, X(t))V(t, X_t) - Du(t, \hat{X}(t))V(t, \hat{X}_t)\Big)\,\mathrm{d}t \\ &+2\gamma\left|\tilde{Z}(t)\right|^{2\gamma-2}\tilde{Z}(t)^\top\Big(Du(t, X(t))\sigma(t, X(t)) - Du(t, \hat{X}(t))\sigma(t, \hat{X}(t))\Big)\,\mathrm{d}W(t) \\ &+\gamma\left|\tilde{Z}(t)\right|^{2\gamma-2}\left\|Du(t, X(t))\sigma(t, X(t)) - Du(t, \hat{X}(t))\sigma(t, \hat{X}(t))\right\|_{HS}^2\mathrm{d}t \\ &+2\gamma(\gamma-1)\left|\tilde{Z}(t)\right|^{2\gamma-4}\left|\Big(Du(t, X(t))\sigma(t, X(t)) - Du(t, \hat{X}(t))\sigma(t, \hat{X}(t))\Big)^\top\tilde{Z}(t)\right|^2\mathrm{d}t.\end{aligned}$$

Using the boundedness of Du and condition (C5) gives for $S \leq t_1 \leq t_2 \leq T$

$$\begin{aligned}&\left|\tilde{Z}(t_2\wedge\tau)\right|^{2\gamma} - \left|\tilde{Z}(t_1\wedge\tau)\right|^{2\gamma} \\ \leq& c\int_{t_1\wedge\tau}^{t_2\wedge\tau}\left\|\tilde{Z}_s\right\|_\infty^{2\gamma}\mathrm{d}s \\ &+c\int_{t_1\wedge\tau}^{t_2\wedge\tau}\left|\tilde{Z}(s)\right|^{2\gamma-1}\left\|Du(s, X(s)) - Du(s, \hat{X}(s))\right\|_{HS}|V(s, X_s)|\,\mathrm{d}s \\ &+c\int_{t_1\wedge\tau}^{t_2\wedge\tau}\left|\tilde{Z}(s)\right|^{2\gamma-2}\tilde{Z}(s)^\top\Big(Du(s, X(s))\sigma(s, X(s)) - Du(s, \hat{X}(s))\sigma(s, \hat{X}(s))\Big)\,\mathrm{d}W(s) \\ &+c\int_{t_1\wedge\tau}^{t_2\wedge\tau}\left|\tilde{Z}(s)\right|^{2\gamma-2}\left\|Du(s, X(s))\sigma(s, X(s)) - Du(s, \hat{X}(s))\sigma(s, \hat{X}(s))\right\|_{HS}^2\mathrm{d}s \\ =&I_1 + I_2 + I_3 + I_4\end{aligned}$$

where $c > 0$ is a constant depending only on $\gamma, d, p, q, \kappa, K, T_0, \|b\|_{L_p^q(T_0)}, g$ and R. The idea is to apply the stochastic Gronwall Lemma A.5. To get rid of the badly behaving terms I_2 and I_4, one can use a suitable multiplier of the form $e^{-A(t)}$ - like in [3] - where

A is an adapted, continuous process. Here, we choose

$$A(t) := c \int_{S\wedge\tau}^{t\wedge\tau} |V(s, X_s)| \frac{\left\| Du(s, X(s)) - Du(s, \hat{X}(s)) \right\|_{HS}}{\left| \tilde{Z}(s) \right|} \mathbb{1}_{\tilde{Z}(s)\neq 0} \mathrm{d}s$$
$$+ c \int_{S\wedge\tau}^{t\wedge\tau} \frac{\left\| Du(s, X(s))\sigma(s, X(s)) - Du(s, \hat{X}(s))\sigma(s, \hat{X}(s)) \right\|_{HS}^2}{\left| \tilde{Z}(s) \right|^2} \mathbb{1}_{\tilde{Z}(s)\neq 0} \mathrm{d}s$$

for $S \wedge \tau \leq t \leq T \wedge \tau$. To show that A is indeed well defined, it suffices to show the existence of a constant $\hat{C} = \hat{C}(\gamma, d, p, q, \kappa, K, T_0, \|b\|_{L_p^q(T_0)}, \|\nabla\sigma\|_{L_p^q(T_0)}, g, R) \geq 0$ such that

$$\mathbb{E} \exp\left(\frac{1}{2} A(T \wedge \tau) \right) \leq \hat{C}.$$

Since u belongs coordinatewise to $H_{2,p}^q(T_0)$ and by conditions (C2) and (C3), it holds

$$(Du \cdot \sigma)^{i,j} \in L^q \left(T_0; W^{1,p} \left(\mathbb{R}^d \right) \right), \ i, j = 1, \ldots, d.$$

Additionally, $C_c^\infty \left(\mathbb{R}^{d+1} \right)$ is dense in $L^q \left(T_0; W^{1,p} \left(\mathbb{R}^d \right) \right)$. Hence, by Young's inequality, Lemmas 2.7 and 2.8, it suffices to show for all $\tilde{R} > 0$ the existence of a constant $C_{\tilde{R}} = C_{\tilde{R}}(d, p, q, \kappa, T_0, \|b\|_{L_p^q(T_0)}, g, R)$ such that

$$\mathbb{E} \exp\left(\int_{S\wedge\tau}^{T\wedge\tau} \frac{\left| f(s, X(s)) - f(s, \hat{X}(s)) \right|^2}{\left| \tilde{Z}(s) \right|^2} \mathbb{1}_{\tilde{Z}(s)\neq 0} \mathrm{d}s \right) \leq C_{\tilde{R}}$$

for all $f \in C^\infty \left(\mathbb{R}^{d+1} \right)$ with $\|f\|_{L^q\left(T_0; W^{1,p}\left(\mathbb{R}^d\right)\right)} \leq \tilde{R}$. By Lemmas A.4 and 2.8, one obtains

$$\begin{aligned} &\mathbb{E} \exp\left(\int_{S\wedge\tau}^{T\wedge\tau} \frac{\left| f(s, X(s)) - f(s, \hat{X}(s)) \right|^2}{\left| \tilde{Z}(s) \right|^2} \mathbb{1}_{\tilde{Z}(s)\neq 0} \mathrm{d}s \right) \\ \leq &\mathbb{E} \exp\left(C_d^2 \int_{S\wedge\tau}^{T\wedge\tau} \left(\mathcal{M} |\nabla f| (X(s)) + \mathcal{M} |\nabla f| (\hat{X}(s)) \right)^2 \mathrm{d}s \right) \\ \leq &C_{\tilde{R}} \end{aligned}$$

where $C_{\tilde{R}} = C_{\tilde{R}}(d, p, q, \kappa, T_0, \|b\|_{L_p^q(T_0)}, g, R)$. By the Itō formula, it holds

$$e^{-A(t)} \left| \tilde{Z}(t \wedge \tau) \right|^{2\gamma} \leq \left| \tilde{Z}(S \wedge \tau) \right|^{2\gamma} + c \int_{S\wedge\tau}^{t\wedge\tau} e^{-A(s)} \left\| \tilde{Z}_s \right\|_\infty^{2\gamma} \mathrm{d}s + \text{local martingale}.$$

Applying the stochastic Gronwall Lemma A.5 gives

$$\mathbb{E} \left[\sup_{S\wedge\tau \leq t \leq T\wedge\tau} e^{-\frac{1}{2}A(t)} \left| \tilde{Z}(t) \right|^\gamma \right] \leq \tilde{C} \mathbb{E} \left\| \tilde{Z}_{S\wedge\tau} \right\|_\infty^\gamma \leq \tilde{C} C_1 \|x - \hat{x}\|_\infty^\gamma$$

for a constant $\tilde{C} = \tilde{C}(\gamma, d, p, q, \kappa, K, T_0, \|b\|_{L_p^q(T_0)}, \|\nabla\sigma\|_{L_p^q(T_0)}, g, R)$. Due to the estimates from above, the Cauchy-Schwarz inequality and by redefining $\gamma := 2\gamma$, one finally obtains

$$\begin{aligned}
&\mathbb{E}\left[\sup_{S\wedge\tau\le t\le T\wedge\tau}\left|\tilde{Z}(t)\right|^{\gamma}\right] \\
\le&\left(\mathbb{E}e^{\frac{1}{2}A(T\wedge\tau)}\right)^{\frac{1}{2}}\left[\mathbb{E}\left(\sup_{S\wedge\tau\le t\le T\wedge\tau} e^{-\frac{1}{2}A(t)}\left|\tilde{Z}(t)\right|^{2\gamma}\right)\right]^{\frac{1}{2}} \\
\le& C_2\|x-\hat{x}\|_{\infty}^{\gamma}
\end{aligned}$$

for some constant $C_2 = C_2(\gamma, d, p, q, \kappa, K, T_0, \|b\|_{L_p^q(T_0)}, \|\nabla\sigma\|_{L_p^q(T_0)}, g, R)$. □

Remark 3.2. The application of the stochastic Gronwall Lemma A.5 is crucial in the proof above, since one has to deal with the supremum norm of path segments. Another standard ansatz might be to apply Doob's or Burkholder's inequality. Unfortunately, it does not work due to the bad regularity of the quadratic variation term of the martingale part. Thus, the inequalities used in [13] or [3] are not suitable.

Proof of Theorem 1.4. This theorem is a consequence of Theorems 2.6, 3.1, Lemma 2.9 and the Yamada-Watanabe Theorem (cf. [12]). □

Proof of Theorem 1.5. Firstly, assume that V is bounded and conditions (C1), (C2) and (C3) are fulfilled. Let $X, \hat{X} \in \mathcal{S}^{\tau}(x)$ for a stopping time τ. Sets of the type

$$\mathcal{K}_n := \left\{x \in \mathcal{C} : \sup_{t\in[-r,0]}|x(t)| + \sup_{-r\le s<t\le 0}\frac{|x(t)-x(s)|}{|t-s|^{1/4}} \le n\right\}$$

with $n \in \mathbb{N}$ are compact in $\mathcal{C}$ and by Lemma 2.9, it holds

$$\lim_{n\to\infty}\tau^n = \tau$$

for

$$\tau^n := \inf\left\{t \le \tau : X_t \notin \mathcal{K}_n \text{ or } \hat{X}_t \notin \mathcal{K}_n\right\} \wedge \tau \wedge n.$$

By assumption, for each $n \in \mathbb{N}$, there exists a $C_{\mathcal{K}_n,n} > 0$ such that V is $C_{\mathcal{K}_n,n}$-Lipschitz continuous in space on $\mathcal{K}_n$. So, there exists a measurable, bounded, in space $C_{\mathcal{K}_n,n}$-Lipschitz continuous extension $V^n : \mathbb{R}_{\ge 0}\times\mathcal{C} \to \mathbb{R}^d$ of $V\big|_{[0,n]\times\mathcal{K}_n}$, i.e.

$$V^n(t,x) = V(t,x)\ \forall x \in \mathcal{K}_n,\ 0 \le t \le n.$$

By Theorem 1.4, for each $n \in \mathbb{N}$, there exists a global, unique strong solution X^n for equation (1) with coefficients V^n, b and σ. Therefore, one has

$$X(t) = \hat{X}(t) = X^n(t),\ 0 \le t \le \tau^n,$$

which provides the pathwise uniqueness.
For the general case, let again $X, \hat{X} \in \mathcal{S}^\tau$ for some stopping time τ. Since V is bounded on compact sets, one has

$$\lim_{n\to\infty} \tau_n = \tau$$

for

$$\tau_n := \inf\left\{t \leq \tau : |V(t, X_t)| > n, |X(t)| > n, \left|V(t, \hat{X}_t)\right| > n \text{ or } \left|\hat{X}(t)\right| > n\right\} \wedge \tau \wedge n.$$

Define for each $n \in \mathbb{N}$

$$b^n(t,x) := \mathbb{1}_{t,|x|\leq n} b(t,x),\ (t,x) \in \mathbb{R}_{\geq 0} \times \mathbb{R}^d$$

and

$$\sigma^n(t,x) := \sigma(t, \phi^n(x)),\ (t,x) \in \mathbb{R}_{\geq 0} \times \mathbb{R}^d$$

where $\phi^n : \mathbb{R}^d \to B_{n+1}$ is a C^1-diffeomorphism defined by

$$\phi(x) := \begin{cases} x, & |x| \leq n, \\ \rho^n(|x|)\frac{x}{|x|}, & |x| > n \end{cases}$$

with

$$\rho^n(t) := n + 1 - \left(\frac{1}{\alpha_n} t - \frac{1}{\alpha_n} n + 1\right)^{-\alpha_n},\ t > n,$$

$$\alpha_n := \frac{p_{n+1} - d}{2}.$$

Furthermore, redefine

$$V^n(t,x) := (-n) \vee V(t,x) \wedge n.$$

By the previous discussion, one has for each $n \in \mathbb{N}$ a global, unique strong solution X^n of equation (1) with coefficients V^n, b^n and σ^n and it holds

$$X(t) = X^n(t) = \hat{X}(t),\ 0 \leq t \leq \tau_n$$

which provides the pathwise uniqueness and the stated maximal solution until the explosion time ζ. □

Proof of Theorem 1.6. The Lipschitz condition (C5) is not necessary for any result in section 2. Accordingly, one can follow the proof of Theorem 3.1 in exactly the same way. Using the same notation, the term

$$\int_{t_1\wedge\tau}^{t_2\wedge\tau} 2\gamma \left|\tilde{Z}(s)\right|^{2\gamma-2} \tilde{Z}(s)^\top \left(Du(s, X(s))V(s, X_s) - Du(s, \hat{X}(s))V(s, \hat{X}_s)\right) \mathrm{d}s$$

was split into

$$\begin{aligned}
&\int_{t_1\wedge\tau}^{t_2\wedge\tau} 2\gamma \left|\tilde{Z}(s)\right|^{2\gamma-2} \tilde{Z}(s)^\top \Big(Du(s,X(s))V(s,X_s) - Du(s,\hat{X}(s))V(s,\hat{X}_s)\Big)\,\mathrm{d}s \\
&= \int_{t_1\wedge\tau}^{t_2\wedge\tau} 2\gamma \left|\tilde{Z}(s)\right|^{2\gamma-2} \tilde{Z}(s)^\top \Big(Du(s,\hat{X}(s))V(s,X_s) - Du(s,\hat{X}(s))V(s,\hat{X}_s)\Big)\,\mathrm{d}s \\
&\quad + \int_{t_1\wedge\tau}^{t_2\wedge\tau} 2\gamma \left|\tilde{Z}(s)\right|^{2\gamma-2} \tilde{Z}(s)^\top \Big(Du(s,X(s))V(s,X_s) - Du(s,\hat{X}(s))V(s,X_s)\Big)\,\mathrm{d}s
\end{aligned}$$

where $S \le t_1 \le t_2 \le T$. The Lipschitz condition (C5) was only used to estimate the first summand by using

$$\left|V(t,X_t) - V(t,\hat{X}_t)\right| \le K \left\|X_t - \hat{X}_t\right\|_\infty,\ t \ge 0.$$

If one can show that the same inequality still holds for two solutions $X \in \mathcal{S}^\tau(x)$ and $\hat{X} \in \mathcal{S}^\tau(x)$ with the same initial value $x \in \mathcal{C}$, the claimed pathwise uniqueness will follow. Now, one has

$$X(t) - \hat{X}(t) = \int_0^t \Big(b(s,X(s)) - b(s,\hat{X}(s))\Big)\,\mathrm{d}s$$

since σ is assumed to be space-independent. Together with Lemma 2.8, it follows that a.s.

$$\Big([-r,t] \ni s \mapsto X(s) - \hat{X}(s)\Big) \in \mathcal{H}_t.$$

Consequently, by rewriting

$$V(t,X_t) - V(\hat{X}_t) = V(t,X_t) - V(t, X_t + \hat{X}_t - X_t),$$

one can apply the assumption and ends up with the desired estimate. The global existence is given by Theorem 2.6. □

A. Appendices

Theorem A.1. *Assume conditions* (C1) *and* (C2)*. Then for any* $T > 0$ *and* $f \in L_p^q(T)$*, there exists a unique solution* $u \in H_{2,p}^q(T)$ *of the following PDE*

$$\begin{aligned}
\partial_t u(t,x) + L_t u(t,x) + f(t,x) &= 0, \\
u(T,x) &= 0
\end{aligned}$$

where

$$L_t u(t,x) := \frac{1}{2}\sum_{i,j,k=1}^d \sigma^{i,k}(t,x)\sigma^{j,k}(t,x)\partial_i\partial_j u(t,x) + b(t,x)\cdot\nabla u(t,x)$$

with the bound

$$\|u\|_{H_{2,p}^q(S,T)} \le C\,\|f\|_{L_p^q(S,T)}$$

for any $S \in [0,T]$ *and some constant* $C = C(T,\kappa,p,q,\|b\|_{L_p^q(T)}) > 0$.

Proof. See [13]. ☐

Theorem A.2. *Let $p, q \in (1, \infty)$, $T > 0$ and $u \in H^q_{2,p}(T)$.*

1. *If $\frac{d}{p} + \frac{2}{q} < 2$, then u is a bounded Hölder continuous function on $[0,T] \times \mathbb{R}^d$ and for any $0 < \varepsilon,\ \delta \leq 1$ satisfying*

$$\varepsilon + \frac{d}{p} + \frac{2}{q} < 2, \quad 2\delta + \frac{d}{p} + \frac{2}{q} < 2,$$

there exists a constant $N = N(p, q, \varepsilon, \delta)$ such that

$$|u(t,x) - u(s,x)| \leq N\,|t-s|^{\delta}\, \|u\|^{1-\frac{1}{q}-\delta}_{\mathbb{H}^q_{2,p}(T)} \|\partial_t u\|^{\frac{1}{q}+\delta}_{L^q_p(T)},$$
$$|u(t,x)| + \frac{|u(t,x) - u(t,y)|}{|x-y|^{\varepsilon}} \leq NT^{-\frac{1}{q}} \left(\|u\|_{\mathbb{H}^q_{2,p}(T)} + T\,\|\partial_t u\|_{L^q_p(T)} \right)$$

for all $s, t \in [0,T]$ and $x, y \in \mathbb{R}^d, x \neq y$.

2. *If $\frac{d}{p} + \frac{2}{q} < 1$, then ∇u is a bounded Hölder continuous function on $[0,T] \times \mathbb{R}^d$ and for any $\varepsilon \in (0,1)$ satisfying*

$$\varepsilon + \frac{d}{p} + \frac{2}{q} < 1,$$

there exists a constant $N = N(p, q, \varepsilon)$ such that

$$|\nabla u(t,x) - \nabla u(s,x)| \leq N\,|t-s|^{\delta}\, \|u\|^{1-\frac{1}{q}-\frac{\varepsilon}{2}}_{\mathbb{H}^q_{2,p}(T)} \|\partial_t u\|^{\frac{1}{q}+\frac{\varepsilon}{2}}_{L^q_p(T)},$$
$$|\nabla u(t,x)| + \frac{|\nabla u(t,x) - \nabla u(t,y)|}{|x-y|^{\varepsilon}} \leq NT^{-\frac{1}{q}} \left(\|u\|_{\mathbb{H}^q_{2,p}(T)} + T\,\|\partial_t u\|_{L^q_p(T)} \right)$$

for all $s, t \in [0,T]$ and $x, y \in \mathbb{R}^d, x \neq y$.

Proof. See [2, p. 22, 23, 36]. ☐

In the next lemma we identify every $u \in H^q_{2,p}$ with its regular version.

Lemma A.3 (Itō formula for $H^q_{2,p}$-functions)**.** *Let $T > 0$, $p > 1$ and $q > 1$ satisfying* (3)*. Let $X : \Omega \times [0,T] \to \mathbb{R}^d$ be a semimartingale on some filtrated probability space $\left(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0}\right)$ of the form*

$$\mathrm{d}X(t) = b(t)\mathrm{d}t + \sigma(t)\mathrm{d}W(t)$$

where $(W_t)_{t\geq 0}$ is a standard d-dimensional Brownian motion, $b : \Omega \times [0,T] \to \mathbb{R}^d$ and $\sigma : \Omega \times [0,T] \to \mathbb{R}^{d\times d}$ are progressively measurable with

$$\mathbb{P}\left(\|b\|_{L^1[0,T]} + \left\|a^{i,j}\right\|_{L^\delta[0,T]} < \infty \right) = 1, \ i, j = 1, \ldots, d$$

for some $1 < \delta \leq \infty$ *where* $a := \sigma\sigma^\top$. *Furthermore, assume that there exists a constant* $C > 0$ *with*

$$\mathbb{E}\int_0^T f(t, X(t))\mathrm{d}t \leq C \|f\|_{L^{q/\delta^*}_{p/\delta^*}(T)}$$

for all $f \in L^{q/\delta^*}_{p/\delta^*}(T)$ *where* δ^* *denotes the conjugate exponent of* δ. *Then for any* $u \in H^q_{2,p}(T)$, *the Itō formula holds, i.e.*

$$\begin{aligned} u(t, X(t)) - u(0, X(0)) = &\int_0^t \partial_t u(s, X(s))\mathrm{d}s + \int_0^t \nabla u(s, X(s))^\top b(s)\mathrm{d}s \\ &+ \int_0^t \nabla u(s, X(s))^\top \sigma(s)\mathrm{d}W(s) \\ &+ \frac{1}{2}\sum_{i,j=1}^d \int_0^t \partial_i\partial_j u(s, X(s))a^{i,j}(s)\mathrm{d}s. \end{aligned}$$

Proof. One can assume without loss of generality

$$\|b\|_{L^1(\Omega\times[0,1])} + \left\|a^{i,j}\right\|_{L^\delta(\Omega\times[0,T])} < \infty,\ i,j = 1, \ldots, d$$

by using the standard localization argument via stopping times. Next, choose a sequence $(u_n)_{n\in\mathbb{N}} \subset C_c^\infty\left(\mathbb{R}^{d+1}\right)$ such that

$$\lim_{n\to\infty} \|u - u_n\|_{H^q_{2,p}(T)} = 0.$$

By the embedding Theorem A.2, it holds

$$\lim_{n\to\infty}\left(\|u_n - u\|_{L^\infty(\mathbb{R}^{d+1})} + \|\nabla u - \nabla u_n\|_{L^\infty(\mathbb{R}^{d+1},\mathbb{R}^d)}\right) = 0.$$

The Itō formula gives for each $n \in \mathbb{N}$ and $t \in [0, T]$

$$\begin{aligned} u_n(t, X(t)) - u_n(0, X(0)) = &\int_0^t \partial_t u_n(s, X(s))\mathrm{d}s + \int_0^t \nabla u_n(s, X(s))^\top b(s)\mathrm{d}s \\ &+ \int_0^t \nabla u_n(s, X(s))^\top \sigma(s)\mathrm{d}W(s) \\ &+ \frac{1}{2}\sum_{i,j=1}^d \int_0^t \partial_i\partial_j u_n(s, X(s))a^{i,j}(s)\mathrm{d}s. \end{aligned}$$

The left-hand side converges to $u(t, X(t)) - u(0, X(0))$ by the choice of u_n. Furthermore, for $\delta < \infty$, one has the following four estimates

$$\begin{aligned}
&\mathbb{E}\int_0^t |\partial_t u(s,X(s)) - \partial_t u_n(s,X(s))| \,\mathrm{d}s \\
\leq &T^{\frac{1}{\delta}} \left(\mathbb{E}\int_0^T |\partial_t u(t,X(t)) - \partial_t u_n(t,X(t))|^{\delta^*} \,\mathrm{d}t\right)^{\frac{1}{\delta^*}} \\
\leq &CT^{\frac{1}{\delta}} \left\| |\partial_t u - \partial_t u_n|^{\delta^*} \right\|_{L^{q/\delta^*}_{p/\delta^*}(T)}^{\frac{1}{\delta^*}} \\
= &CT^{\frac{1}{\delta}} \left\| \partial_t u - \partial_t u_n \right\|_{L^q_p(T)},
\end{aligned}$$

$$\mathbb{E}\int_0^t |\nabla u(s) - \nabla u_n(s)|\, |b(s)| \,\mathrm{d}s \leq \|\nabla u - \nabla u_n\|_{L^\infty(\mathbb{R}^{d+1},\mathbb{R}^d)} \|b\|_{L^1(\Omega\times[0,T])},$$

$$\begin{aligned}
&\mathbb{E}\left|\int_0^t (\nabla u(s,X(s)) - \nabla u_n(s,X(s)))^\top \sigma(s)\mathrm{d}W(s)\right|^2 \\
\leq &\mathbb{E}\int_0^T |(\nabla u(t,X(t)) - \nabla u_n(t,X(t)))\,\sigma(t)|^2 \,\mathrm{d}t \\
\leq &\|\nabla u - \nabla u_n\|^2_{L^\infty(\mathbb{R}^{d+1},\mathbb{R}^d)} \cdot \mathbb{E}\int_0^T \|\sigma(t)\|^2_{HS} \,\mathrm{d}t,
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}\sum_{i,j=1}^d \int_0^t |\partial_i\partial_j u(s,X(s)) - \partial_i\partial_j u_n(s,X(s))| \left|a^{i,j}(s)\right| \mathrm{d}s \\
\leq &\sum_{i,j=1}^d \left(\mathbb{E}\int_0^T |\partial_i\partial_j u(t,X(t)) - \partial_i\partial_j u_n(t,X(t))|^{\delta^*} \,\mathrm{d}t\right)^{\frac{1}{\delta^*}} \left\|a^{i,j}\right\|_{L^\delta(\Omega\times[0,T])} \\
\leq &\sum_{i,j=1}^d C \left\| |\partial_i\partial_j u - \partial_i\partial_j u_n|^{\delta^*} \right\|_{L^{q/\delta^*}_{p/\delta^*}(T)}^{\frac{1}{\delta^*}} \left\|a^{i,j}\right\|_{L^\delta(\Omega\times[0,T])} \\
= &\sum_{i,j=1}^d C \left\| \partial_i\partial_j u - \partial_i\partial_j u_n \right\|_{L^q_p(T)} \left\|a^{i,j}\right\|_{L^\delta(\Omega\times[0,T])}.
\end{aligned}$$

For $\delta = \infty$, the estimates are basically the same. All these terms above converge to zero by the choice of u_n, which provides the desired convergence of the right-hand side. $\square$

Let ϕ be a locally integrable function on $\mathbb{R}^d$. The Hardy-Littlewood maximal function is definded by

$$\mathcal{M}\phi(x) := \sup_{0<r<\infty} \frac{1}{|B_r|} \int_{B_r} \phi(x+y)\mathrm{d}y$$

where B_r is the Euclidean ball of radius r. The following result is cited from [13].

Lemma A.4.

1. *There exists a constant $C_d > 0$ such that for all $\phi \in C^\infty(\mathbb{R}^d)$ and $x, y \in \mathbb{R}^d$,*
$$|\phi(x) - \phi(y)| \leq C_d |x - y| (\mathcal{M} |\nabla\phi| (x) + \mathcal{M} |\nabla\phi| (y)).$$
2. *For any $p > 1$, there exists a constant $C_{d,p}$ such that for all $\phi \in L^p(\mathbb{R}^d)$,*
$$\|\mathcal{M}\phi\|_{L^p} \leq C_{d,p} \|\phi\|_{L^p}.$$

Lemma A.5. *Let Z be an adapted non-negative stochastic process with continuous paths defined on $[0, \infty)$ which satisfies the inequality*
$$Z(t) \leq K \int_0^t \sup_{0 \leq r \leq s} Z(r) \mathrm{d}s + M(t) + C,$$
where $C \geq 0$, $K \geq 0$ and M is a continuous local martingale with $M(0) = 0$. Then for each $0 < p < 1$, there exist universal finite constants $c_1(p)$, $c_2(p)$ (not depending on K, C, T and M) such that
$$\mathbb{E}\left[\sup_{0 \leq t \leq T} Z(t)^p\right] \leq C^p c_2(p) e^{c_1(p)KT} \text{ for every } T \geq 0.$$

Proof. See [11, 9]. □

Lemma A.6 (Modified Khas'minskii lemma)**.** *Let $\beta : \Omega \times [0, T] \to \mathbb{R}_{\geq 0}$ be a non-negative, measurable, adapted process with respect to some filtrated probability space $\left(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{0 \leq t \leq T}\right)$ and $T > 0$. Assume there exists some $0 \leq \alpha < 1$ for all $0 \leq s \leq t \leq T$ such that*
$$\mathbb{E}\left[\int_s^t \beta(r) \mathrm{d}r \middle| \mathcal{F}_s\right] \leq \alpha.$$
Then one has
$$\mathbb{E} \exp\left(\int_0^T \beta(r) \mathrm{d}r\right) \leq \frac{1}{1 - \alpha}.$$

Proof. This proof mainly follows the technique used in [2].

$$
\begin{aligned}
&\mathbb{E}\exp\left(\int_0^T \beta(r)\mathrm{d}r\right)\\
&= \sum_{n=0}^{\infty} \frac{1}{n!}\mathbb{E}\int_0^T \cdots \int_0^T \beta(r_1)\cdots\beta(r_n)\mathrm{d}r_1\cdots\mathrm{d}r_n\\
&= \sum_{n=0}^{\infty} \mathbb{E}\int_{0\le r_1\le r_2\le\cdots\le r_n\le T} \beta(r_1)\cdots\beta(r_n)\mathrm{d}r\\
&= \sum_{n=0}^{\infty} \mathbb{E}\left[\int_{0\le r_1\le r_2\le\cdots\le r_{n-1}\le T} \beta(r_1)\cdots\beta(r_{n-1})\int_{r_{n-1}}^T \beta(r_n)\mathrm{d}r_n\mathrm{d}r_1\cdots\mathrm{d}r_{n-1}\right]\\
&= \sum_{n=0}^{\infty} \mathbb{E}\left[\int_0^T \cdots \int_0^{r_2} \beta(r_1)\cdots\beta(r_{n-1})\mathbb{E}\left(\int_{r_{n-1}}^T \beta(r_n)\mathrm{d}r_n\middle|\mathcal{F}_{r_{n-1}}\right)\mathrm{d}r_1\cdots\mathrm{d}r_{n-1}\right]\\
&\le \sum_{n=0}^{\infty} \alpha\mathbb{E}\int_0^T \cdots \int_0^{r_2} \beta(r_1)\cdots\beta(r_{n-1})\mathrm{d}r_1\cdots\mathrm{d}r_{n-1}\\
&\overset{(\star)}{\le} \sum_{n=0}^{\infty} \alpha^n\\
&= \frac{1}{1-\alpha}
\end{aligned}
$$

where $(\star)$ was obtained by iteration. □

References

[1] O. Enchev and D. W. Stroock. Rademacher's theorem for Wiener functionals. *Ann. Probab.*, 21(1):25–33, 1993.

[2] E. Fedrizzi. Uniqueness and flow theorems for solutions of SDEs with low regularity of the drift. *tesi di Laurea in Matematica, Università di Pisa*, 2009.

[3] E. Fedrizzi and F. Flandoli. Pathwise uniqueness and continuous dependence of SDEs with non-regular drift. *Stochastics*, 83(3):241–257, 2011.

[4] F. Flandoli, M. Gubinelli, and E. Priola. Flow of diffeomorphisms for SDEs with unbounded Hölder continuous drift. *Bull. Sci. Math.*, 134(4):405–422, 2010.

[5] I. Gyöngy and T. Martínez. On stochastic differential equations with locally unbounded drift. *Czechoslovak Math. J.*, 51(126)(4):763–783, 2001.

[6] N. V. Krylov and M. Röckner. Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Related Fields*, 131(2):154–196, 2005.

[7] R. S. Liptser and A. N. Shiryaev. *Statistics of Random Processes: I. General Theory*, pages 286–297. Springer, 2001.

[8] N. I. Portenko. *Generalized diffusion processes*, volume 83 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1990. Translated from the Russian by H. H. McFaden.

[9] M. Scheutzow. A stochastic gronwall lemma. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 16(02):1350019, 2013.

[10] A. Ju. Veretennikov. Strong solutions of stochastic differential equations. *Teor. Veroyatnost. i Primenen.*, 24(2):348–360, 1979.

[11] M.-K. von Renesse and M. Scheutzow. Existence and uniqueness of solutions of stochastic functional differential equations. *Random Oper. Stoch. Equ.*, 18(3):267–284, 2010.

[12] T. Yamada and S. Watanabe. On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.*, 11(1):155–167, 1971.

[13] X. Zhang. Stochastic homeomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients. *Electron. J. Probab.*, 16:no. 38, 1096–1116, 2011.

[14] X. Zhang. Stochastic differential equations with Sobolev drifts and driven by α-stable processes. *Ann. Inst. Henri Poincaré Probab. Stat.*, 49(4):1057–1079, 2013.

[15] A. K. Zvonkin. A transformation of the phase space of a diffusion process that will remove the drift. *Mat. Sb. (N.S.)*, 93(135):129–149, 152, 1974.

On the Strong Feller Property for Stochastic Delay Differential Equations with Singular Drift

Stefan Bachmann

Email: bachmann@math.uni-leipzig.de

Institut für Mathematik, Universität Leipzig,
Augustusplatz 10, 04109 Leipzig, Germany

December 8, 2019

Abstract

In this paper, we prove the strong Feller property for stochastic delay (or functional) differential equations with singular drift. We extend an approach of Maslowski and Seidler to derive the strong Feller property of those equations, see [14]. The argumentation is based on the well-posedness and the strong Feller property of the equations' drift-free version. To this aim, we investigate a certain convergence of random variables in topological spaces in order to deal with discontinuous drift coefficients.

Keywords: Stochastic delay differential equations, stochastic functional differential equation, strong Feller property, singular drift, Zvonkin's transformation.

MSC2010: Primary 34K50; secondary 60B10, 60B12, 60H10.

1. Introduction

In this paper, we investigate a certain convergence of random variables in topological spaces and apply the results to prove an improved version of the strong Feller property of the following stochastic delay (or functional) differential equation (SDDE).

$$\begin{aligned} \mathrm{d}X^x(t) &= B(t, X_t^x)\mathrm{d}t + b(t, X^x(t))\mathrm{d}t + \sigma(t, X^x(t))\mathrm{d}W(t), \\ X_0^x &= x \end{aligned} \tag{1}$$

where W is a d-dimensional Brownian motion, $B : \mathbb{R}_{\geq 0} \times C\left([-r, 0], \mathbb{R}^d\right) \to \mathbb{R}^d$ is measurable and strictly sublinear in the second variable, $\sigma : \mathbb{R}_{\geq 0} \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ is measurable, bounded, non-degenerate and Lipschitz in space and $b \in L^q_{loc}\left(\mathbb{R}_{\geq 0}; L^p\left(\mathbb{R}^d\right)\right)$ where $1 < p, q$ fulfill

$$\frac{d}{p} + \frac{2}{q} < 1. \tag{2}$$

The well-posedness and stability of equation (1) without a delay drift term, i.e. $B \equiv 0$, has been studied thoroughly: Krylov and Röckner have shown existence and uniqueness for the special case $\sigma \equiv Id$ in [12], which had elaborated previous results, for example of Portenko [15], Veretennikov [22] and Zvonkin [27]. In [17], Rutkowski considered pathwise uniqueness for stochastic one-dimensional equations with singular drift involving local time conditions which were introduced by Barlow and Perkins. Martínez and Gyöngy have proven well-posedness for non-constant σ in [10], however, under the stricter assumption $b \in L^{2d+2}(\mathbb{R}^{d+1})$. In [26], Zhang was able to improve it further by only assuming $\left|\nabla_x \sigma^{i,j}\right| \in L^q_{loc}\left(\mathbb{R}_{\geq 0}; L^p\left(\mathbb{R}^d\right)\right)$, $i, j = 1, \ldots, d$. Gawarecki and Mandrekar have studied SDEs with discontinuous drift and their connection to unbounded spin-systems in [9]. Blei and Engelbert have analyzed one-dimensional equations with so-called generalized drifts in [2] and haven given sufficient and necessary criteria for existence and uniqueness in distribution. A nonlinear Kolmogorov equation for stochastic functional delay differential equations with jumps has been proven by Cordoni, Di Persio and Oliva in [4]. Cordoni and Di Persio have also shown existence and uniqueness of mild solutions for stochastic reaction-diffusion equations on networks with dynamic time-delayed boundary conditions in [3] and have considered an application to a stochastic optimal control problem. In [1], we extended the well-posedness result of Zhang [26] to the delay case, essentially by using a combination of Zvonkin's transformation, several Girsanov techniques and a stochastic Gronwall lemma from von Renesse and Scheutzow in [18, 23]. The existence of a unique strong solution for SDDEs with singular drift has also been shown in [11].

Da Prato, Elworthy and Zabczyk have studied the strong Feller property for stochastic semilinear equations with unbounded coefficients in [16]. The strong Feller property with respect to the state space $\mathbb{R}^d$ of equation (1) with unbounded non-functional drift has been shown by Zhang in [26]. However, in this paper we are interested in the state space of path segments $C([-r, 0], \mathbb{R}^d)$. Several Harnack inequalities have been studied for stochastic delay differential equations, which imply the strong Feller property. Es-Sarhir, von Renesse and Scheutzow have investigated the case $b \equiv 0$ and $\sigma \equiv const$ in [6]. Wang and Yuan have established results for non-constant and uniformly non-degenerate diffusion coefficients, which do not depend on the past, in [24]. By remark 1.4 in [6], the strong Feller property might not be given if the diffusion term is of real functional nature. Both papers are based on a coupling technique.

However, we do not use a coupling technique but the probabilistic approach of Maslowski and Seidler, see [14]. In order to deal with discontinuous drift coefficients, we consider a certain convergence for topological spaces, see Theorem 1.8. As a result, we gain a simple method to derive the strong Feller property of equation (1) from the well-posedness and the strong Feller property of the simpler special case $B \equiv 0$ and $b \equiv 0$.

Notation 1.1. We denote by $\|\cdot\|_{OP}$ and $\|\cdot\|_{HS}$ the operator norm and respectively the Hilbert-Schmidt norm for matrices $A \in \mathbb{R}^{d\times d}$, i.e.

$$\|A\|_{op} = \sup_{v\in\mathbb{R}^d, |v|=1} |Av|\,,\ \|A\|_{HS} = \sqrt{\sum_{i,j=1}^{d} |A^{i,j}|^2}.$$

Additionally, we write for $a, b \in [-\infty, +\infty]$

$$a \wedge b := \min\{a, b\},\ a \vee b := \max\{a, b\}.$$

Notation 1.2. In the sequel, let $r > 0$ be an arbitrary but fixed number and define

$$\mathcal{C} := C\left([-r, 0], \mathbb{R}^d\right)$$

equipped with the supremum norm $\|\cdot\|_\infty$. For a process X defined on $[t-r, t]$ with $t \geq 0$, we write

$$X_t(s) := X(t+s),\ s \in [-r, 0].$$

Furthermore, we introduce the following function spaces, which will be used later on: define for $0 \leq S \leq T < \infty$ and $p, q \in (1, \infty)$

$$\begin{aligned}
L_p^q(S,T) &:= L^q\left([S,T]; L^p\left(\mathbb{R}^d\right)\right), & L_p^q(T) &:= L_p^q(0,T),\\
\mathbb{H}_{2,p}^q(S,T) &:= L^q\left([S,T]; W^{2,p}\left(\mathbb{R}^d\right)\right), & \mathbb{H}_{2,p}^q(T) &:= \mathbb{H}_{2,p}^q(0,T),\\
H_{2,p}^q(S,T) &:= W^{1,q}\left([S,T]; L^p\left(\mathbb{R}^d\right)\right) \cap \mathbb{H}_{2,p}^q(,T), & H_{2,p}^q(T) &:= H_{2,p}^q(0,T),
\end{aligned}$$

equipped with the norm

$$\|u\|_{H_{2,p}^q(S,T)} := \|\partial_t u\|_{L_p^q(S,T)} + \|u\|_{\mathbb{H}_{2,p}^q(S,T)}\,,\ u \in H_{2,p}^q(S,T).$$

Notation 1.3. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be some probability space. Then we denote by $\mathbb{P}^*$ the corresponding outer measure, i.e.

$$\mathbb{P}^*(A) := \inf\left\{\mathbb{P}(B) : A \subseteq B, B \in \mathcal{F}\right\},\ A \subseteq \Omega.$$

Notation 1.4. In the sequel, we always equip topological spaces with their corresponding Borel σ-algebra, and subspaces with the usual subspace topology. Furthermore, if $(E, \mathcal{E})$ is some measurable space, we denote by $B_b(E)$ the space of bounded, measurable functions.

Notation 1.5. If not stated otherwise, W will be a d-dimensional Brownian motion on some arbitrary but fixed probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and every strong solution shall be defined on this space.

However, weak solutions of equation (1) might be defined on different filtrated probability spaces. Therefore, we use the short hand notation $(X^x, \tilde{W}^x, \mathbb{Q}^x)$ where X^x is an adapted, continuous stochastic process, $\tilde{W}^x$ is an adapted Brownian motion, both with respect to some filtrated probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mathbb{Q}^x, (\tilde{\mathcal{F}}_t)_{t\geq 0})$, and $(X^x, \tilde{W}^x)$ solves equation (1) with initial value x.

Condition C1. *Let $p, q > 1$ be given with* (2)*. One has for every $T > 0$*

$$b \in L_p^q(T).$$

Condition C2. *Assume that for all $T > 0$ there exists some $C_\sigma = C_\sigma(T) > 0$ such that*

1. $C_\sigma^{-1} I_{d\times d} \le \sigma(t,x)\sigma(t,x)^\top \le C_\sigma I_{d\times d}\ \forall t \in [0,T], x \in \mathbb{R}^d$,
2. $\|\sigma(t,x) - \sigma(t,y)\|_{HS} \le C_\sigma |x-y|\ \ \forall t \in [0,T], x, y \in \mathbb{R}^d$.

Condition C3. *For $t \in [0,r)$ the function $x \mapsto B(t,x)$ is continuous and for all $T > 0$ there exists some monotone increasing $g_T : \mathbb{R}_{\ge 0} \to \mathbb{R}_{\ge 0}$ with*

1. $|B(t,x)| \le g_T(\|x\|_\infty)\ \forall x \in \mathcal{C}, t \in [0,T]$,
2. $\lim\limits_{r\to\infty} g_T(r)/r = 0$.

Condition C4. *For all $T > 0$ there exists some $C_B = C_B(T) > 0$ such that*

$$|B(t,x) - B(t,y)| \le C_B \|x-y\|_\infty\ \ \forall t \in [0,T], x, y \in \mathcal{C}.$$

The main results read as follows.

Theorem 1.6. *Assume conditions* (C1), (C2) *and* (C3)*. Then for each initial value $x \in \mathcal{C}$, equation* (1) *has a global weak solution $(X^x, \tilde{W}^x, \mathbb{Q}^x)$, which is unique in distribution. Furthermore, one has the strong Feller property for all $t > r$, i.e.*

$$\lim_{y\to x} \mathbb{E}_{\mathbb{Q}^y} f(X_t^y) = \mathbb{E}_{\mathbb{Q}^x} f(X_t^x)\ \forall f \in B_b(\mathcal{C}).$$

Moreover, if condition (C4) *is fulfilled, then equation* (1) *has a unique strong solution and it holds*

$$\lim_{y\to x} \mathbb{E}_{\mathbb{P}} |f(X_t^y) - f(X_t^x)| = 0\ \forall f \in B_b(\mathcal{C}).$$

The following theorem is the key element for proving Theorem 1.6.

Theorem 1.7. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be some probability space and (E, d) be a metric space. Furthermore, let $X, X_n : \Omega \to E$, $n \in \mathbb{N}$ be measurable maps. Then the statement*

1. *a)* $\lim\limits_{n\to\infty} \mathbb{P}^*\left(d\left(X, X_n\right) \ge \varepsilon\right) = 0\ \forall \varepsilon > 0$,

   b) $\lim\limits_{n\to\infty} \mathbb{P}_{X_n}(O) = \mathbb{P}_X(O)$, *for all open $O \subset E$*

implies

2. $\lim\limits_{n\to\infty} \mathbb{E}\,|f(X) - f(X_n)| = 0\ \forall f \in B_b(\mathcal{C})$.

Additionally, if there exists some null set $N \subset \Omega$ such that $X(\Omega \setminus N)$ is separable, then the converse implication is also true.

Moreover, we give a version of Theorem 1.7 in a topologically more general setup, which reads as follows (the topological terminologies are given in subsection 2.1).

Theorem 1.8. *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be some probability space and* $(E, \mathcal{D})$ *be a gauge space. Furthermore, let* $X, X_n : \Omega \to E$*,* $n \in \mathbb{N}$ *be measurable maps and assume that* $\mathbb{P}_X$ *is outer regular and there exists some null set* $N \subset \Omega$ *such that* $X(\Omega \setminus N)$ *is Lindeöf. Then the following statements are equivalent*

1. *a)* $\lim_{n\to\infty} \mathbb{P}^* \left(d\left(X, X_n\right) \geq \varepsilon\right) = 0 \ \forall \varepsilon > 0, d \in \mathcal{D}$,

   b) $\lim_{n\to\infty} \mathbb{P}_{X_n}(O) = \mathbb{P}_X(O)$ *for all open* $O \subset E$.

2. $\lim_{n\to\infty} \mathbb{E}\left|f(X) - f(X_n)\right| = 0 \ \forall f \in B_b(E)$.

Remark 1.9. To the best of our knowledge, this kind of convergence has not been studied systematically. In this work, it is a key element for proving the strong Feller property together with the approach of Maslowski and Seidler, see [14]. In subsection 3.1, we introduce the underlying strategy, which is applicable to a much more general setup.

Additionally, in subsection 2.3 we give some examples to compare the different convergence concepts.

Remark 1.10. The continuity assumption in condition (C3) might look artificial. However, the following example illustrates that the strong Feller property is not given in general if one drops this assumption. Consider the SDDE (with $r = 1$)

$$\begin{aligned} \mathrm{d}X^x(t) &= \operatorname{sgn}\left(X^x(t-1)\right)\mathrm{d}t + \mathrm{d}W(t), \\ X_0^x &= x \end{aligned}$$

with the convention

$$\operatorname{sgn} x = \begin{cases} 1 & \text{if } x \geq 0, \\ -1 & \text{if } x < 0. \end{cases}$$

This equation has for each initial value a unique strong solution, which can be constructed recursively. Now, set $y_n \equiv -1/n$, $n \in \mathbb{N}$, then one has

$$\begin{aligned} X^0(1) &= W(1) + 1, \\ X^{y_n}(1) &= W(1) - 1 - \frac{1}{n}. \end{aligned}$$

Thus, the strong Feller property is not given.

Remark 1.11. In appendix B, we consider the strict topology on the space of bounded, continuous functions as an example of a non-metrizable, locally convex space where all assumptions of Theorem 1.8 are fulfilled. The strict topology is used for Markov processes with a state space that is not locally compact, see [20].

2. Convergence of Random Variables in Topological Spaces and Examples

2.1. Preliminaries

Definition 2.1. A topological space X is called Lindelöf if every open cover of X has a countable subcover. X is called hereditarily Lindelöf if every open set of X is Lindelöf with respect to the subspace topology.

Definition 2.2. Let E be some nonempty set and $\mathcal{D}$ be a nonempty set of pseudometrics on E. Then we call $(E, \mathcal{D})$ a gauge space and its topology shall be generated by

$$\left\{B_r^d(x) : x \in E, d \in \mathcal{D}, r > 0\right\}$$

where

$$B_r^d(x) := \{y \in E : d(x,y) < r\}, \ x \in E, d \in \mathcal{D}, r > 0.$$

Definition 2.3. A Borel probability measure $\mathbb{P}$ on a topological space $(E, \mathcal{T})$ is called outer regular if for all Borel sets $A \in \mathcal{B}(E)$ and $\varepsilon > 0$ there exists some open $O \in \mathcal{T}$ with $A \subseteq O$ such that

$$\mathbb{P}(O \setminus A) < \varepsilon.$$

2.2. Convergence of Random Variables in Topological Spaces

At the beginning we prove the following abstract lemma.

Lemma 2.4. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be some probability space and $(E, \mathfrak{E})$ be a measurable space. Furthermore, let $X_n : \Omega \to E$, $n \in \mathbb{N}$ be a sequence of measurable maps, $X : \Omega \to E$ be measurable and $\mathfrak{S} \subseteq \mathfrak{E}$ such that*

$$\forall A \in \mathfrak{E}, \varepsilon > 0 \ \exists S \in \mathfrak{S} : \mathbb{P}_X(A \Delta S) < \varepsilon$$

where Δ denotes the symmetric difference of sets. Then the following statements are equivalent

1. *a)* $\lim_{n\to\infty} \mathbb{P}\left(X \in S, X_n \notin S\right) = 0$ *for all* $S \in \mathfrak{S}$,

   b) $\lim_{n\to\infty} \mathbb{P}_{X_n}(A) = \mathbb{P}_X(A)$ *for all* $A \in \mathfrak{E}$.

2. $\lim_{n\to\infty} \mathbb{E}\left|f(X) - f(X_n)\right| = 0 \ \forall f \in B_b(E)$.

Proof. Implication 2. ⇒ 1. is trivial. Hence, we only show implication 1. ⇒ 2. Since f is bounded, it suffices to prove for all $A \in \mathfrak{E}$

$$\lim_{n\to\infty} \mathbb{P}\left(X \in A, X_n \notin A\right) = 0.$$

Let $\varepsilon > 0$. By assumption, there exists some $S \in \mathfrak{S}$ such that

$$\mathbb{P}_X(A \Delta S) < \varepsilon.$$

Now, choose $n_0 \in \mathbb{N}$ large enough such that

$$\mathbb{P}_{X_n}(A \Delta S) < \varepsilon$$

for all $n \geq n_0$. Then one has for all $n \geq n_0$

$$\mathbb{P}(X \in A, X_n \notin A) \leq \mathbb{P}(X \in S, X_n \notin S) + 2\varepsilon.$$

By assumption, one obtains

$$\lim_{n\to\infty} \mathbb{P}(X \in A, X_n \notin A) \leq 2\varepsilon.$$

Since $\varepsilon > 0$ was chosen arbitrarily, the proof is complete. $\square$

Remark 2.5. Let $(E, \mathcal{T})$ be a topological space and $(\mathbb{P}_n)_{n\in\mathbb{N}}$ be a sequence of probability measures that converges pointwise on $\mathcal{T}$ to some outer regular probability measure, i.e.

$$\lim_{n\to\infty} \mathbb{P}_n(O) = \mathbb{P}(O) \ \forall O \in \mathcal{T}.$$

Then one has pointwise convergence on all Borel sets:

$$\lim_{n\to\infty} \mathbb{P}_n(A) = \mathbb{P}(A) \ \forall A \in \mathcal{B}(E).$$

This can be seen as follows. Let $A \in \mathcal{B}(E)$ and $\varepsilon > 0$. Since $\mathbb{P}$ was assumed to be outer regular, one can find an open set $O \supseteq A$ and a closed set $C \subseteq A$ such that

$$\mathbb{P}(O \setminus C) < \varepsilon.$$

By assumption, it follows

$$\mathbb{P}(C) = \lim_{n\to\infty} \mathbb{P}_n(C) \leq \liminf_{n\to\infty} \mathbb{P}_n(A) \leq \limsup_{n\to\infty} \mathbb{P}_n(A) \leq \lim_{n\to\infty} \mathbb{P}_n(O) = \mathbb{P}(O)$$

and consequently,

$$\mathbb{P}(A) - \varepsilon \leq \liminf_{n\to\infty} \mathbb{P}_n(A) \leq \limsup_{n\to\infty} \mathbb{P}_n(A) \leq \mathbb{P}(A) + \varepsilon \ \forall \varepsilon > 0.$$

Now, one can apply the previous lemma to the topological context, which reads as follows.

Corollary 2.6. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be some probability space and $(E, \mathcal{T})$ be a topological space. Furthermore, let $X_n : \Omega \to E$, $n \in \mathbb{N}$ be a sequence of measurable maps and $X : \Omega \to E$ be measurable such that the measure $\mathbb{P}_X$ is outer regular. Then the following statements are equivalent*

1. *for all open $O \in \mathcal{T}$ holds*
   a) $\lim_{n\to\infty} \mathbb{P}(X \in O, X_n \notin O) = 0$,
   b) $\lim_{n\to\infty} \mathbb{P}_{X_n}(O) = \mathbb{P}_X(O)$.

2. $\lim_{n\to\infty} \mathbb{E}\left|f(X) - f(X_n)\right| = 0 \ \forall f \in B_b(E)$.

Remark 2.7. For a separable metric space (E, d), it holds

$$\mathcal{B}(E \times E) = \mathcal{B}(E) \otimes \mathcal{B}(E).$$

If one drops the separability assumption, this may fail. See for example $E = 2^{\mathbb{R}}$ equipped with the discrete topology. Consequently, for two measurable random variables $X, Y : \Omega \to E$, the map

$$\Omega \ni \omega \mapsto d(X(\omega), Y(\omega))$$

could be not measurable. To overcome this problem, one can use the outer measure

$$\mathbb{P}^*(A) := \inf\{\mathbb{P}(B) : A \subset B, \ B \text{ measurable}\}$$

to evaluate

$$\mathbb{P}^*(d(X, Y) \geq \varepsilon), \ \varepsilon > 0.$$

This provides a natural definition of convergence in probability for non-separable metric spaces. For a discussion in detail, see [21].

Lemma 2.8. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be some probability space and $(E, \mathcal{D})$ be a gauge space. Furthermore, let $X, X_n : \Omega \to E$, $n \in \mathbb{N}$ be measurable maps and assume that $\mathbb{P}_X$ is outer regular and there exists some null set $N \subset \Omega$ such that $X(\Omega \setminus N)$ is Lindelöf. Then the following statements are equivalent*

1. $\lim_{n\to\infty} \mathbb{P}^*(d(X, X_n) \geq \varepsilon) = 0 \ \forall \varepsilon > 0, d \in \mathcal{D}$,
2. $\lim_{n\to\infty} \mathbb{P}(X \in O, X_n \notin O) = 0$ *for all open* $O \subset E$.

Proof. 1. $\Rightarrow$ 2. : Without loss of generality, one can assume

$$\max(d_1, d_2) \in \mathcal{D} \ \forall d_1, d_2 \in \mathcal{D}.$$

Let $O \subseteq E$ be open and $\delta > 0$. The probability measure $\mathbb{P}_X$ was assumed to be outer regular. Thus, there exists an open set V with $E \setminus O \subseteq V$ and

$$\mathbb{P}_X(V \cap O) < \delta.$$

Additionally, for every $x \in O$ there exist some pseudometric $d_x \in \mathcal{D}$ and $r_x > 0$ such that

$$O = \bigcup_{x \in O} B_{r_x}^{d_x}(x).$$

Then it follows

$$E = \bigcup_{x \in O} B_{r_x}^{d_x}(x) \cup V,$$

and since $X(\Omega \setminus N)$ is Lindelöf, one can find $d_i \in \mathcal{D}$, $r_i > 0$, $x_i \in E$, $i \in \mathbb{N}$ such that

$$\mathbb{P}_X \left(O \setminus \bigcup_{i=1}^{m} B_{r_i}^{d_i}(x_i) \right) < \delta.$$

Then it holds

$$\begin{aligned}
\mathbb{P}\left(X \in O, X_n \notin O\right) &\leq \sum_{i=1}^{m} \mathbb{P}\left(X \in B_{r_i}^{d_i}(x_i), X_n \notin O\right) + \delta \\
&\leq \sum_{i=1}^{m} \mathbb{P}\left(X \in B_{r_i}^{d_i}(x_i), X_n \notin B_{r_i}^{d_i}(x_i)\right) + \delta
\end{aligned}$$

Hence, it suffices to show

$$\lim_{n\to\infty} \mathbb{P}\left(X \in B_r^d(x), X_n \notin B_r^d(x)\right) = 0 \ \forall d \in \mathcal{D}, x \in E, r > 0.$$

Let $d \in \mathcal{D}$, $x \in E$ and $r > 0$. By assumption, one has

$$\lim_{n\to\infty} \mathbb{P}^*\left(d\left(X, X_n\right) \geq \varepsilon\right) = 0 \ \forall \varepsilon > 0.$$

Now, we show that every subsequence of $(X_n)_{n\in\mathbb{N}}$ has a subsequence that converges pointwise to X with respect to d almost surely. This can be seen as follows. Let $(X_{n_k})_{k\in\mathbb{N}}$ be a subsequence of $(X_n)_{n\in\mathbb{N}}$. Now, for each $l \in \mathbb{N}$ choose $k_l \in \mathbb{N}$ large enough such that $k_l \geq k_{l-1}$ and

$$\mathbb{P}^*\left(d\left(X, X_{k_l}\right) \geq \frac{1}{l}\right) \leq \frac{1}{l^2}.$$

This is possible by the equality above. Then one has by the properties of an outer measure

$$\begin{aligned}
&\mathbb{P}^*\left(X_{k_l} \nrightarrow X\right) \\
\leq &\mathbb{P}^*\left(\bigcup_{m\in\mathbb{N}} \bigcap_{l_0\in\mathbb{N}} \bigcup_{l\geq l_0} \left\{d\left(X, X_{n_{k_l}}\right) \geq \frac{1}{m}\right\}\right) \\
\leq &\sum_{m\in\mathbb{N}} \mathbb{P}^*\left(\bigcap_{l_0\in\mathbb{N}} \bigcup_{l\geq l_0} \left\{d\left(X, X_{n_{k_l}}\right) \geq \frac{1}{m}\right\}\right).
\end{aligned}$$

Furthermore, one has

$$
\begin{aligned}
&\mathbb{P}^*\left(\bigcap_{l_0\in\mathbb{N}}\bigcup_{l\geq l_0}\left\{d\left(X,X_{n_{k_l}}\right)\geq\frac{1}{m}\right\}\right)\\
=&\mathbb{P}^*\left(\bigcap_{l_0\in\mathbb{N}}\bigcup_{l\geq l_0}\left\{d\left(X,X_{n_{k_l}}\right)\geq\frac{1}{l}\right\}\right)\\
\leq&\lim_{l_0\to\infty}\sum_{l=l_0}^{\infty}\mathbb{P}^*\left(d\left(X,X_{k_l}\right)\geq\frac{1}{l}\right)\\
\leq&\lim_{l_0\to\infty}\sum_{l=l_0}^{\infty}\frac{1}{l^2}\\
=&0.
\end{aligned}
$$

Consequently, every subsequence of $(X_n)_{n\in\mathbb{N}}$ has a subsequence that converges pointwise to X with respect to d almost surely. Finally, it follows

$$\lim_{n\to\infty}\mathbb{P}\left(X\in B_r^d(x),X_n\notin B_r^d(x)\right)=0.$$

2. $\Rightarrow$ 1. : Let ε, $\delta>0$ and $d\in\mathcal{D}$. Since $X(\Omega\setminus N)$ was assumed to be Lindelöf, one can choose $m\in\mathbb{N}$ large enough such that

$$\mathbb{P}_X\left(\bigcup_{i=1}^{m}B_\varepsilon^d(x_i)\right)>1-\delta$$

with suitable $x_1,\ldots,x_m\in E$. Then one has

$$
\begin{aligned}
&\mathbb{P}^*\left(d\left(X,X_n\right)\geq\varepsilon\right)\\
\leq&\mathbb{P}^*\left(X\in\bigcup_{i=1}^{m}B_\varepsilon^d(x),d(X,X_n)\geq\varepsilon\right)+\mathbb{P}^*\left(X\notin\bigcup_{i=1}^{m}B_\varepsilon^d(x_i),d(X,X_n)\geq\varepsilon\right)\\
\leq&\sum_{i=1}^{m}\mathbb{P}^*\left(X\in B_\varepsilon^d(x_i),d\left(X,X_n\right)\geq\varepsilon\right)+\mathbb{P}\left(X\notin\bigcup_{i=1}^{m}B_\varepsilon^d(x_i)\right)\\
\leq&\sum_{i=1}^{m}\mathbb{P}\left(X\in B_\varepsilon^d(x_i),X_n\notin B_\varepsilon^d(x_i)\right)+\delta
\end{aligned}
$$

By assumption, it follows

$$\lim_{n\to\infty}\mathbb{P}^*\left(d\left(X,X_n\right)\geq\varepsilon\right)\leq\delta$$

for all $\delta>0$, which completes the proof. $\square$

Definition 2.9. In a topological space a G_δ-set is an intersection of countably many open sets. A topological space is called a G_δ-space if every closed set is a G_δ-set.

Lemma 2.10. *Let $(E, \mathcal{T})$ be a G_δ-space. Then every Borel probability measure $\mathbb{P}$ on E is outer regular.*

Proof. The proof is standard and can be found for polish spaces in [5, p. 224-225]. Consider the set $\mathfrak{A}$

$$\mathfrak{A} := \{A \in \mathcal{B}(E) : A \text{ and } E \setminus A \text{ outer regular}\}.$$

Clearly, $\mathfrak{A}$ is a σ-algebra. Therefore, it is sufficient to show that every closed $C \subseteq E$ is contained in $\mathfrak{A}$. By assumption, each closed set is a countable intersection of open sets, which completes the proof. □

Remark 2.11. Examples for G_δ-spaces are metric spaces and hereditarily Lindelöf gauge spaces.

Proof of Theorem 1.7. By Corollary 2.6 and Lemma 2.8, it is sufficient to show

$$\lim_{n\to\infty} \mathbb{P}\left(X \in O, X_n \notin O\right) = 0 \ \forall \text{ open } O \subset E.$$

Furthermore,

$$\lim_{n\to\infty} \mathbb{P}^*\left(d(X, X_n) \leq \varepsilon\right) = 0 \ \forall \varepsilon > 0$$

implies that every subsequence of $(X_n)_{n\in\mathbb{N}}$ has a subsequence that converges almost surely. Thus, it follows

$$\lim_{n\to\infty} \mathbb{1}_O(X)\mathbb{1}_{E\setminus O}(X_n) = 0 \text{ in probability},$$

which proofs the implication. Under the additional assumption, one can apply Lemma 2.8 again to conclude the reverse direction. □

Proof of Theorem 1.8. This is a consequence of Corollary 2.6, Lemma 2.8 and Lemma 2.10. □

2.3. Examples

In this subsection we want to give some examples for the reader's convenience. At first, two rather trivial examples are given to discuss the equivalence from Theorem 1.7. Then we consider two one-dimensional SD(D)Es to investigate the difference between the strong Feller property and its "improved" version.

1. In this simple example we show that pointwise convergence of random variables does not imply the convergence type discussed in this section. However, if one adds Gaussian terms, the convergence follows. Consider a sequence $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}^d$ that converges to some $x_0 \in \mathbb{R}^d$ with $x_i \neq x_0$ for all $i \in \mathbb{N}$. Then the deterministic random variables

$$X_n \equiv x_n$$

converge pointwise to x_0 but their laws δ_{x_n} do not converge pointwise to δ_{x_0}. In particular, it holds

$$\mathbb{E}\left|\mathbb{1}_{\{x_0\}}(X) - \mathbb{1}_{\{x_0\}}(X_n)\right| = 1.$$

On the other hand, let N be a standard Gaussian random variable and consider instead the sequence

$$Y_n := x_n + N.$$

Then one has

$$\lim_{n\to\infty} \mathbb{E}\left|f(Y_n) - f(Y)\right| = 0 \ \forall f \in B_b(\mathbb{R}).$$

2. This example is rather trivial. It shows that pointwise convergence of measures does not imply the convergence type discussed in this section. Let $X \sim \mathcal{N}(0,1)$ be a standard Gaussian random variable and define

$$X_n := -X \sim \mathcal{N}(0,1).$$

It holds

$$\mathbb{P}(X \in A) = \mathbb{P}(X_n \in A) \ \forall n \in \mathbb{N}, A \in \mathcal{B}(\mathbb{R}).$$

Obviously, X_n does not converge to X in probability and it holds

$$\mathbb{E}\left|\mathbb{1}_{\mathbb{R}_{\geq 0}}(X) - \mathbb{1}_{\mathbb{R}_{\geq 0}}(X_n)\right| = 1 \ \forall n \in \mathbb{N}.$$

3. In this example we show that pathwise uniqueness and the strong Feller property already implies almost sure convergence for solutions of one-dimensional non-delay equations. Consequently, in this setting, there is no difference between the strong Feller property and it's "improved" version. Now, assume we have a one-dimensional SDE that has a unique strong solution for each real initial value

$$\begin{aligned} \mathrm{d}X^x(t) &= b(t, X^x(t))\mathrm{d}t + \sigma(t, X^x(t))\mathrm{d}W(t), \\ X^x(0) &= x \in \mathbb{R} \end{aligned}$$

where W is a d-dimensional Brownian motion on some probability space, and $b : \mathbb{R}_{\geq 0} \times \mathbb{R} \to \mathbb{R}$ and $\sigma : \mathbb{R}_{\geq 0} \times \mathbb{R} \to \mathbb{R}^{1,d}$ are measurable. Assume furthermore that X has the Feller property, i.e.

$$\lim_{y\to x} \mathbb{E}f(X^y(t)) = \mathbb{E}f(X^x(t)) \ \forall f \in C_b(\mathbb{R}).$$

Then for every sequence $(x_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ with $x_n \to x$ and $t > 0$, one has

$$X^{x_n}(t) \to X^x(t) \text{ a.s.}$$

In particular, by Theorem 1.7, the strong Feller property is equivalent to

$$\lim_{n\to\infty} \mathbb{E}\left|f(X^{x_n}(t)) - f(X^x(t))\right| = 0 \ \forall f \in B_b(\mathbb{R}).$$

This can be seen as follows: by uniqueness, one has monotonicity for the solutions, i.e. for all $x \le y$ holds

$$X^x(t) \le X^y(t) \ \forall t \ge 0 \text{ a.s.}$$

On the other hand, for each sequence $(x_n)_{n\in\mathbb{N}}$ with $x_n \downarrow x$, the following limit exists

$$\tilde{X}(t) := \lim_{x_n \downarrow x} X^{x_n}(t) \ \forall t \ge 0 \text{ a.s.}$$

since all X^{x_n} are bounded from below by X^x. Thus, one has for all $t \ge 0$, $f \in C_b(\mathbb{R})$

$$\mathbb{E}f(X^x(t)) = \lim_{x_n \downarrow x} \mathbb{E}f(X^{x_n}(t)) = \mathbb{E}f(\tilde{X}(t)).$$

Additionally, it holds

$$\tilde{X}(t) \ge X^x(t) \ \forall t \ge 0 \text{ a.s.}$$

Thus, $\tilde{X}$ is a modification of X^x.

4. In this example we show that the implication from the previous example is not given if the dispersion coefficient depends on the past. Especially, condition (C2) is not fulfilled. For the sake of overview, we embed real constants in $\mathcal{C}$ naturally. Now, let us consider the one-dimensional SDDE (with $r = 1$)

$$\begin{aligned} \mathrm{d}X^x(t) &= \operatorname{sgn}(X^x(t-1))\mathrm{d}W(t) \\ X_0 &= x \in \mathcal{C} \end{aligned}$$

where we use the convention

$$\operatorname{sgn} x = \begin{cases} 1 & \text{if } x \ge 0, \\ -1 & \text{if } x < 0. \end{cases}$$

This SDDE can be solved uniquely by constructing the solution recursively. By Levy's characterization, each solution X^x is distributed on $\mathbb{R}_{\ge 0}$ like a shifted Brownian motion, in particular

$$X_t^x \sim W_t + x(0) \ \forall t > 1.$$

It is not difficult to show that one has for $t > 1$

$$\lim_{y\to x} \mathbb{E}f(X_t^y) = \lim_{y\to x} \mathbb{E}f(W_t + y(0)) = \mathbb{E}f(W_t + x(0)) = \mathbb{E}f(X_t^x) \ \forall f \in B_b(\mathbb{R}),$$

see for example [6]. So, X has the strong Feller property with respect to the state space $\mathcal{C}$. On the other hand, one has for all $y \ge 0, x < 0$

$$\|X_2^y - X_2^x\|_\infty \ge |X^y(1) - X^x(1)| = |2W(1) + y - x| \text{ a.s.}$$

Therefore, convergence in probability is not given. In particular, the strong Feller property does not coincide with it's "improved" version.

3. Application to Stochastic Delay Differential Equations

3.1. Introduction to the Method

In this subsection, we want to illustrate the strategy to prove Theorem 1.6, which is based on an approach of Maslowski and Seidler, see [14], and the convergence discussed in section 2. In order, we make use of a toy example with state space $\mathbb{R}^d$. Consider the equation

$$\begin{aligned}\mathrm{d}X^x(t) &= b(t, X^x(t))\mathrm{d}t + \sigma \mathrm{d}W(t),\\ X^x(0) &= x \in \mathbb{R}^d\end{aligned}$$

where W is some d-dimensional Brownian motion, $\sigma \in \mathbb{R}^{d\times d}$ is invertible, $b : \mathbb{R}_{\geq 0} \times \mathbb{R}^d \to \mathbb{R}^d$ is measurable and for every $T > 0$ there exists some $C_T \in \mathbb{R}$ with

$$\begin{aligned}\langle b(t,u) - b(t,v), u - v\rangle \leq C_T \left|u - v\right|^2 \ \ \forall u, v \in \mathbb{R}^d, t \in [0, T],\\ |b(t,u)| \leq C_T \ \forall u \in \mathbb{R}^d, t \in [0, T].\end{aligned}$$

Additionally, we consider its drift-free equation

$$\begin{aligned}\mathrm{d}M^x(t) &= \sigma \mathrm{d}W(t),\\ M^x(0) &= x \in \mathbb{R}^d,\end{aligned}$$

which is trivial in that case for the sake of simplicity. Observe that both equations have a unique strong solution and the drift-free one depends continuously on the initial value in the sense that

$$\lim_{n\to\infty} M^{x_n}(t) = M^x(t) \text{ in probability}$$

for each sequence $(x_n)_{n\in\mathbb{N}} \subset \mathbb{R}^d$ converging to x. Now, one can observe that M has the strong Feller property, i.e.

$$\lim_{y\to x} \mathbb{E}f(M^y(t)) = \mathbb{E}f(M^x(t)) \ \forall f \in B_b(\mathbb{R}^d).$$

Also, $\mathbb{P}_{X^x}$ has a Girsanov density with respect to $\mathbb{P}_{M^x}$, i.e.

$$\mathbb{E}f(X^x(t)) = \mathbb{E}\left[D^x(t)f(M^x(t))\right] \ \ \forall f \in B_b(\mathbb{R}^d)$$

with

$$D^x(t) = \exp\left(\int_0^t \sigma^{-1}b(s, M^x(s))^\top \mathrm{d}W(s) - \frac{1}{2}\int_0^t \left|\sigma^{-1}b(s, M^x(s))\right|^2 \mathrm{d}s\right).$$

At first, we show that X has the strong Feller property, too. Let $f \in B_b\left(\mathbb{R}^d\right)$, then one has

$$\begin{aligned}&\mathbb{E}f(X^x(t)) - \mathbb{E}f(X^y(t))\\ =&\mathbb{E}[D^x(t)f(M^x(t))] - \mathbb{E}[D^y(t)f(M^y(t))]\\ \leq&\mathbb{E}[D^x(t)(f(M^x(t)) - f(M^y(t)))] + \|f\|_\infty \mathbb{E}\left|D^x(t) - D^y(t)\right|.\end{aligned}$$

By Theorem 1.7, one has

$$\lim_{y \to x} \mathbb{E}\, |f(M^x(t)) - f(M^y(t))| = 0$$

and in particular,

$$\lim_{n \to \infty} D^x(t) f(M^{x_n}(t)) = D^x(t) f(M^x(t)) \text{ in probability}$$

for each sequence $(x_n)_{n \in \mathbb{N}} \subset \mathbb{R}^d$ converging to x. By the dominated convergence theorem, it follows

$$\lim_{y \to x} \mathbb{E}[D^x(t)(f(M^y(t)) - f(M^x(t)))] = 0.$$

Consequently, it remains to show

$$\lim_{y \to x} \mathbb{E}\, |D^x(t) - D^y(t)| = 0.$$

Assume for a moment that one has

$$\lim_{n \to \infty} D^{x_n}(t) = D^x(t) \text{ in probability}$$

for each sequence $(x_n)_{n \in \mathbb{N}} \subset \mathcal{C}$ converging to x. Since one has $\mathbb{E}_{\mathbb{P}} D^z(t) = 1$ and $D^z(t) \geq 0$ for all $z \in \mathcal{C}$, one could apply Fatou's lemma to conclude

$$2 - \lim_{n \to \infty} \mathbb{E}_{\mathbb{P}}\, |D^{x_n}(t) - D^x(t)| = \lim_{n \to \infty} \mathbb{E}_{\mathbb{P}}\, (D^x(t) + D^{x_n}(t) - |D^{x_n}(t) - D^x(t)|) \geq 2$$

and the desired L^1-convergence follows. Thus, it suffices to show

$$\lim_{n \to \infty} D^{x_n}(t) = D^x(t) \text{ in probability}$$

for each sequence $(x_n)_{n \in \mathbb{N}} \subset \mathcal{C}$ converging to x. Therefore, it is sufficient to show

$$\lim_{y \to x} \mathbb{E} \int_0^t |b(s, M^y(s)) - b(s, M^x(s))|^2 \, \mathrm{d}s = 0,$$

by the martingale isometry. However, this is a direct consequence of Theorem 1.7. Finally, one ends up with

$$\lim_{y \to x} \mathbb{E} f(X^y(t)) = \mathbb{E} f(X^x(t)).$$

By Itō's formula and Gronwall's lemma, one can easily show

$$\lim_{n \to \infty} X^{x_n}(t) = X^x(t) \text{ in probability}$$

for each sequence $(x_n)_{n \in \mathbb{N}} \subset \mathbb{R}^d$ converging to x. Thus, it even follows

$$\lim_{y \to x} \mathbb{E}\, |f(X^y)(t) - f(X^x(t))| = 0$$

by Theorem 1.7.

The well-posedness and the strong Feller property of the drift-free equation, the existence of densities of X^x with respect to M^x, $x \in \mathbb{R}^d$, and their convergence in probability are exactly the requirements Maslowski and Seidler needed for one of their approaches to show the strong Feller property, cf. Theorem 2.1 in [14]. Theorem 1.7 systematically extends their approach by showing the convergence of the densities even for discontinuous drift coefficients.

Now, we can summarize the strategy for showing the strong Feller property in a few steps without specifying the details: show that

1. the drift-free version of the original equation has a unique strong solution M^x for each initial value x.
2. M has the strong Feller property and for every sequence $(x_n)_{n\in\mathbb{N}}$ with $x_n \to x$, one has
$$\lim_{n\to\infty} M_t^{x_n} = M_t^x \text{ in probability}$$
3. The equation with drift has for each initial value x a weak solution X^x that is unique in distribution.
4. For every initial value x, $\mathbb{P}_{X^x}$ has a density with respect to $\mathbb{P}_{M^x}$ such that for any sequence $(x_n)_{n\in\mathbb{N}}$ with $x_n \to x$, one has
$$\lim_{x_n\to x} \frac{\mathrm{d}\mathbb{P}_{X^{x_n}}}{\mathrm{d}\mathbb{P}_{M^{x_n}}}(M^{x_n}) = \frac{\mathrm{d}\mathbb{P}_{X^x}}{\mathrm{d}\mathbb{P}_{M^x}}(M^x) \text{ in probability.}$$
As illustrated by the previous example, Theorem 1.7 is the key element for verifying this step since one has to deal with discontinuous coefficients.

If, in addition, the equation with drift has for each initial value x a unique strong solution X^x such that for every sequence $(x_n)_{n\in\mathbb{N}}$ with $x_n \to x$, one has
$$\lim_{n\to\infty} X_t^{x_n} = X_t^x \text{ in probability,}$$
then one can apply Theorem 1.7 again to deduce the "improved" version of the Feller property.

3.2. A-priori Estimates, Uniqueness and Existence

In the sequel, denote by M^x, $x \in \mathcal{C}$ the global, unique strong solution of
$$\begin{aligned} \mathrm{d}M^x(t) &= \sigma(t, M^x(t))\,\mathrm{d}W(t), \\ M_0^x &= x. \end{aligned}$$

Remark 3.1. Condition (C3) implies the following important property of the delay drift B: for all $\alpha > 0$ and $T > 0$ there exists a $K_{\alpha,T} > 0$ such that
$$|B(t,x)| \le \alpha \|x\| + K_{\alpha,T} \;\forall t \in [0,T].$$

In the sequel, this property will be exploited to show that the Novikov-condition is fulfilled for various Girsanov densities.

Moreover, condition (C2) implies the following inequalities

$$\|\sigma\|_{op}, \left\|\sigma^{-1}\right\|_{op} \leq \sqrt{C_\sigma}.$$

Lemma 3.2. *Assume condition* (C2)*. Let $T > 0$ and $p', q' > 1$ be given with*

$$\frac{d}{p'} + \frac{2}{q'} < 2.$$

Then one has for all $0 \leq S < T$ and $f \in L^{q'}_{p'}(S,T)$ the estimate

$$\mathbb{E}\left(\int_S^T f(t, M^x(t))\mathrm{d}t \middle| \mathcal{F}_S\right) \leq C\,\|f\|_{L^{q'}_{p'}(S,T)}$$

for some constant $C = C(d, p', q', T, C_\sigma)$. In particular, the constant C is independent of the initial value $x \in \mathcal{C}$.

Proof. This follows directly from Theorem 2.1 in [26]. □

Lemma 3.3. *Assume condition* (C2)*. Then for any $R, T > 0$ and $p', q' > 1$ with*

$$\frac{d}{p'} + \frac{2}{q'} < 2$$

there exists a constant $C_R = C_R(d, p', q', T, C_\sigma)$ such that

$$\mathbb{E}\exp\left(\int_0^T f(t, M^x(t))\mathrm{d}t\right) \leq C_R$$

for all $f \in L^{q'}_{p'}(T)$ with $\|f\|_{L^{q'}_{p'}(T)} \leq R$.

Proof. See Lemma 2.2 in [1]. □

Lemma 3.4. *Assume condition* (C2)*. Then for any $T > 0$ and $0 \leq \alpha < (2dC_\sigma T)^{-1}$, it holds*

$$\mathbb{E}\exp\left(\alpha \sup_{0\leq t\leq T} |M^x(t)|^2\right) \leq \frac{4}{\sqrt{1-2\alpha d C_\sigma T}}\exp\left(\frac{\alpha}{1-2\alpha d C_\sigma T}|x(0)|^2\right).$$

Proof. See Lemma 2.4 in [1]. □

Theorem 3.5. *Assume conditions* (C1), (C2) *and* (C3)*. Then for every initial values $x \in \mathcal{C}$, equation* (1) *has a global weak solution. Moreover, for each weak solution $(X^x, \tilde{W}^x, \mathbb{Q}^x)$ of equation* (1) *on some time interval $[-r, T]$, $T > 0$, one has*

$$\mathbb{Q}^x_{X^x}(A) = \mathbb{E}_{\mathbb{P}}\left[\mathbb{1}_A(M^x)\exp\left(\int_0^T a^x(t)^\top \mathrm{d}W(t) - \frac{1}{2}\int_0^T |a^x(t)|^2\,\mathrm{d}t\right)\right],$$
$$a^x(t) := \sigma(t, M^x(t))^{-1}\left[B(t, M^x_t) + b(t, M^x(t))\right],\ t \in [0,T]$$

for all measurable $A \subset C([-r,T], \mathbb{R}^d)$. In addition, if condition (C4) *holds, equation* (1) *admits a unique strong solution.*

Proof. At first, we show the existence of a weak solution. The additional statement about the pathwise uniqueness has been shown in [1], Theorem 1.5, and the existence of a strong solution then follows from the Theorem of Yamada and Watanabe [25].

The strong solution M^x is by definition $(\mathcal{F}_t)_{t\geq 0}$-adapted where $(\mathcal{F}_t)_{t\geq 0}$ is the augmented filtration generated by W. Next, we construct a probability measure on

$$\mathcal{F}_\infty := \sigma\left(\mathcal{F}_t : t \geq 0\right)$$

such that M^x is a global weak solution for equation (1). By Lemma 3.3, Lemma 3.4, condition (C2) and condition (C3), it holds

$$\mathbb{E}_{\mathbb{P}} \exp\left(\alpha \int_0^T \left|\sigma(t, M^x(t))^{-1} B(t, M_t^x)\right|^2 \mathrm{d}t\right) < \infty,$$
$$\mathbb{E}_{\mathbb{P}} \exp\left(\alpha \int_0^T \left|\sigma(t, M^x(t))^{-1} b(t, M^x(t))\right|^2 \mathrm{d}t\right) < \infty$$

for all $\alpha \in \mathbb{R}$ and $T > 0$. Therefore, Novikov's condition is fulfilled and Girsanov's theorem is applicable, which gives that

$$\bar{W}(t) := W(t) - \int_0^t \sigma(s, M^x(s))^{-1} a^x(s)\mathrm{d}s,\ t \geq 0$$

is a Brownian motion on $[0, T]$ under the probability measure

$$\begin{aligned}\mathrm{d}\bar{\mathbb{P}}_T := \exp\Bigg(&\int_0^T \left(\sigma(t, M^x(t))^{-1} a^x(t)\right)^\top \mathrm{d}W(t) \\ &- \frac{1}{2}\int_0^T \left|\sigma(t, M^x(t)) a^x(t)\right|^2 \mathrm{d}t\Bigg)\mathrm{d}\mathbb{P}\end{aligned}$$

and $(M^x, \bar{W}, \bar{\mathbb{P}}_T)$ is a weak solution of (1) on $[-r, T]$ for each $T > 0$. Additionally, one has for $0 < T_1 < T_2$

$$\bar{\mathbb{P}}_{T_1}(A) = \bar{\mathbb{P}}_{T_2}(A)\ \forall A \in \mathcal{F}_{T_1},$$

so the probability measure on $\mathcal{F}_\infty$ uniquely defined by

$$\bar{\mathbb{P}}(A) := \mathbb{P}_T(A)\ \forall T > 0, A \in \mathcal{F}_T$$

is indeed well-defined and $(M^x, \bar{W}, \bar{\mathbb{P}})$ is a global weak solution.

Now, let $(X^x, \tilde{W}^x, \mathbb{Q}^x)$ be a weak solution on some time interval $[0, T]$, $T > 0$. The following approach is inspired by the techniques used in [13]. Define

$$\tau^n(\omega) := \inf\left\{s \geq 0 : |\omega(s)| \geq n\right\} \wedge T,\ \omega \in C([-r, T], \mathbb{R}^d),\ n \in \mathbb{N}.$$

Then the stopped process $X^{x,n}(t) := X^x(t \wedge \tau^n(X^x))$, $t \in [-r, T]$ fulfills the equation

$$\mathrm{d}X^{x,n}(t) = \mathbb{1}_{\tau^n(X^{x,n}) \leq t}\left[B(t, X_t^{x,n}) + b(t, X^{x,n}(t))\right]\mathrm{d}t + \mathbb{1}_{\tau^n(X^{x,n}) \leq t}\sigma(t, X^{x,n}(t))\mathrm{d}\tilde{W}^x$$

By condition (C3) and Girsanov's theorem,

$$\tilde{W}^{x,n}(t) := \int_0^{t\wedge\tau^n(X^{x,n})} \sigma(s, X^{x,n}(s))^{-1} B(s, X_s^{x,n})\mathrm{d}s + \tilde{W}^x(t),\ t \geq 0$$

is a Brownian motion with respect to the probability measure

$$\begin{aligned}\mathrm{d}\mathbb{Q}^{x,n} := \exp\Bigg(&- \int_0^{\tau^n(X^{x,n})} \left(\sigma(t, X^{x,n}(t))^{-1} B(t, X_t^{x,n})\right)^\top \mathrm{d}\tilde{W}^x(t) \\ &- \frac{1}{2}\int_0^{\tau^n(X^{x,n})} |\sigma(t, X^{x,n}(t))B(t, X_t^{x,n})|^2\,\mathrm{d}t\Bigg)\mathrm{d}\mathbb{Q}^x.\end{aligned}$$

The process $X^{x,n}$ solves the equation

$$\begin{aligned}\mathrm{d}X^{x,n}(t) &= b(t, X^{x,n}(t))\mathrm{d}t + \sigma(t, X^{x,n}(t))\mathrm{d}\tilde{W}^{x,n}(t),\ t \in [0, \tau^n(X^{x,n})], \\ X_0^{x,n} &= x.\end{aligned}$$

Such a solution is unique by Theorem 1.3 in [26], i.e.

$$X^{x,n}(t) = Y^{x,n}(t),\ t \in [-r, \tau^n(X^{x,n})]$$

where $Y^{x,n}$ is the unique strong solution of

$$\begin{aligned}\mathrm{d}Y^{x,n}(t) &= b(t, Y^{x,n}(t))\mathrm{d}t + \sigma(t, Y^{x,n}(t))\mathrm{d}\tilde{W}^{x,n}(t), \\ Y_0^{x,n} &= x\end{aligned}$$

and it holds

$$\tau^n(X^{x,n}) = \tau^n(Y^{x,n}) \text{ a.s.}$$

Consequently, the process $X^{x,n}$ is distributed, with respect to $\mathbb{Q}^{x,n}$, as the stopped process $t \mapsto Y^{x,n}(t \wedge \tau^n(Y^{x,n}))$, $t \geq -r$. It follows

$$\begin{aligned}&\mathbb{Q}^x(X^x \in A) \\ &= \lim_{n\to\infty} \mathbb{Q}^x(\tau^n(X^x) = T, X^x \in A) \\ &= \lim_{n\to\infty} \mathbb{Q}^x(\tau^n(X^{x,n}) = T, X^{x,n} \in A) \\ &= \lim_{n\to\infty} \mathbb{E}_{\mathbb{Q}^{x,n}}\Bigg[\mathbb{1}_{\tau^n(X^{x,n})=T}\mathbb{1}_A(X^{x,n}) \exp\Bigg(\int_0^T \left(\sigma(t, X^{x,n}(t))^{-1} B(t, X_t^{x,n})\right)^\top \mathrm{d}\tilde{W}^{x,n}(t) \\ &\quad - \frac{1}{2}\int_0^T \left|\sigma(t, X^{x,n}(t))^{-1} B(t, X_t^{x,n})\right|^2 \mathrm{d}t\Bigg)\Bigg] \\ &= \lim_{n\to\infty} \mathbb{E}_{\mathbb{Q}^{x,n}}\Bigg[\mathbb{1}_{\tau^n(Y^{x,n})=T}\mathbb{1}_A(Y^{x,n}) \exp\Bigg(\int_0^T \left(\sigma(t, Y^{x,n}(t))^{-1} B(t, Y_t^{x,n})\right)^\top \mathrm{d}\tilde{W}^{x,n}(t) \\ &\quad - \frac{1}{2}\int_0^T \left|\sigma(t, Y^{x,n}(t))^{-1} B(t, Y_t^{x,n})\right|^2 \mathrm{d}t\Bigg)\Bigg]\end{aligned}$$

for all measurable $A \subset C([-r,T],\mathbb{R}^d)$. On the other hand, by Lemma 3.3, the process

$$\hat{W}(t) := W(t) - \int_0^t \sigma(s, M^x(s))^{-1} b(s, M^x(s))\mathrm{d}s,\ t \geq 0$$

is a Brownian motion under the probability measure

$$\begin{aligned}\mathrm{d}\hat{\mathbb{P}} :=\exp\Bigg(&\int_0^T \left(\sigma(t, M^x(t))^{-1} b(t, M_t^x)\right)^\top \mathrm{d}W(t)\\ &- \frac{1}{2}\int_0^T |\sigma(t, M^x(t))b(t, M_t^x)|^2\,\mathrm{d}t\Bigg)\mathrm{d}\mathbb{P}.\end{aligned}$$

and M^x solves the equation

$$\begin{aligned}\mathrm{d}M^x(t) &= b(t, M^x(t))\mathrm{d}t + \sigma(t, M^x(t))\hat{W}(t)\\ M_0^x &= x.\end{aligned}$$

Again, by uniqueness in distribution, one has $\hat{\mathbb{P}}_{M^x} = \mathbb{Q}^{x,n}_{Y^{x,n}}$ for all $n \in \mathbb{N}$. With this in hand, one obtains

$$\begin{aligned}&\mathbb{E}_{\mathbb{P}}\left[\mathbb{1}_A(M^x)\exp\left(\int_0^T a^x(t)^\top \mathrm{d}W(t) - \frac{1}{2}\int_0^T |a^x(t)|^2\,\mathrm{d}t\right)\right]\\ &= \lim_{n\to\infty}\mathbb{E}_{\mathbb{P}}\left[\mathbb{1}_{\tau^n(M^x)=T}\mathbb{1}_A(M^x)\exp\left(\int_0^T a^x(t)^\top \mathrm{d}W(t) - \frac{1}{2}\int_0^T |a^x(t)|^2\,\mathrm{d}t\right)\right]\\ &= \lim_{n\to\infty}\mathbb{E}_{\hat{\mathbb{P}}}\Bigg[\mathbb{1}_{\tau^n(M^x)=T}\mathbb{1}_A(M^x)\\ &\quad\times\exp\Bigg(\int_0^T B(t, M_t^x)^\top\left(\sigma(t, M^x(t))\sigma(t, M^x(t))^\top\right)^{-1}\mathrm{d}M^x(t)\\ &\quad - \frac{1}{2}\int_0^T (2b(t, M^x(t)) + B(t, M_t^x))^\top\left(\sigma(t, M^x(t))\sigma(t, M^x(t))^\top\right)^{-1} B(t, M_t^x)\mathrm{d}t\Bigg)\Bigg]\\ &= \lim_{n\to\infty}\mathbb{E}_{\mathbb{Q}^{x,n}}\Bigg[\mathbb{1}_{\tau^n(Y^{x,n})=T}\mathbb{1}_A(Y^{x,n})\\ &\quad\times\exp\Bigg(\int_0^T B(t, Y_t^{x,n})^\top\left(\sigma(t, Y^{x,n}(t))\sigma(t, Y^{x,n}(t))^\top\right)^{-1}\mathrm{d}Y^{x,n}(t)\\ &\quad - \frac{1}{2}\int_0^T (2b(t, Y^{x,n}(t)) + B(t, Y_t^{x,n}))^\top\left(\sigma(t, Y^{x,n}(t))\sigma(t, Y^{x,n}(t))^\top\right)^{-1} B(t, Y_t^{x,n})\mathrm{d}t\Bigg)\Bigg]\\ &= \lim_{n\to\infty}\mathbb{E}_{\mathbb{Q}^{x,n}}\Bigg[\mathbb{1}_{\tau^n(Y^{x,n})=T}\mathbb{1}_A(Y^{x,n})\exp\Bigg(\int_0^T \left(\sigma(t, Y^{x,n}(t))^{-1}B(t, Y_t^{x,n})\right)^\top\mathrm{d}\tilde{W}^{x,n}(t)\\ &\quad - \frac{1}{2}\int_0^T \left|\sigma(t, Y^{x,n}(t))^{-1}B(t, Y_t^{x,n})\right|^2\mathrm{d}t\Bigg)\Bigg]\\ &= \mathbb{Q}^x(X^x \in A)\end{aligned}$$

for all measurable $A \subset C([-r,T],\mathbb{R}^d)$. □

Lemma 3.6. *Assume conditions* (C1), (C2) *and* (C3). *Let* $T>0$, $R>0$ *and* $p', q' \in (1,\infty)$ *be given with*

$$\frac{d}{p'} + \frac{2}{q'} < 2.$$

Then for each weak solution $(X^x, \tilde{W}^x, \mathbb{Q}^x)$ *of equation* (1) *on* $[-r,T]$ *with* $\|x\|_\infty \le R$, *one has*

$$\sup_{f \in L^{q'}_{p'}(T) : \|f\|_{L^{q'}_{p'}(T)} \le R} \mathbb{E}_{\mathbb{Q}^x} \exp\left(\int_0^T f(t, X^x(t)) \mathrm{d}t\right) \le C_R.$$

with a constant $C_R = C_R(p,q,p',q',d,T,C_\sigma, \|b\|_{L^q_p(T)}, g_T)$. *Additionally, one has*

$$\mathbb{E}_{\mathbb{Q}^x} \int_0^T f(s, X^x(s)) \mathrm{d}s \le C \|f\|_{L^{q'}_{p'}(T)}$$

with a constant $C>0$.

Proof. As before, let

$$a^x(t) := \sigma(t, M^x(t))^{-1} \left(B(t, M^x_t) + b(t, M^x(t))\right), \ t \in [0,T].$$

By Theorem 3.5, one has

$$\begin{aligned}
&\mathbb{E}_{\mathbb{Q}^x} \exp\left(\int_0^T f(t, X^x(t)) \mathrm{d}t\right) \\
=&\mathbb{E}_{\mathbb{P}} \exp\left(\int_0^T f(t, M^x(t)) \mathrm{d}t + \int_0^T a^x(t)^\top \mathrm{d}W(t) - \frac{1}{2}\int_0^T |a^x(t)|^2 \mathrm{d}t\right) \\
\le& \left[\mathbb{E}_{\mathbb{P}} \exp\left(\int_0^T 2f(t, M^x(t)) \mathrm{d}t\right)\right]^{\frac{1}{2}} \left[\mathbb{E}_{\mathbb{P}} \exp\left(2\int_0^T a^x(t)^\top \mathrm{d}W(t) - \int_0^T |a^x(t)|^2 \mathrm{d}t\right)\right]^{\frac{1}{2}} \\
\le& \left[\mathbb{E}_{\mathbb{P}} \exp\left(\int_0^T 2f(t, M^x(t)) \mathrm{d}t\right)\right]^{\frac{1}{2}} \cdot \left[\mathbb{E}_{\mathbb{P}} \exp\left(6\int_0^T |a^x(t)|^2 \mathrm{d}t\right)\right]^{\frac{1}{4}}.
\end{aligned}$$

The uniform bound follows from condition (C3), Lemma 3.3 and Lemma 3.4. By Lemma 3.2, one has

$$\int_0^T f(t, M^x(t)) \mathrm{d}s \to 0 \text{ in probability}$$

if $\|f\|_{L^{q'}_{p'}(T)} \to 0$. Together with the exponential bound from above, it follows

$$\mathbb{E}_{\mathbb{Q}^x} \int_0^T f(t, X^x(t)) \mathrm{d}t \to 0$$

if $\|f\|_{L^{q'}_{p'}(T)} \to 0$. Consequently, the linear operator $A : L^{q'}_{p'}(T) \to \mathbb{R}$ given by

$$f \mapsto \mathbb{E}_{\mathbb{Q}^x} \int_0^T f(t, X^x(t)) \mathrm{d}t$$

is continuous, which provides the existence of the desired constant. □

Lemma 3.7. *Assume conditions* (C1), (C2) *and* (C3) *and let* $T > 0$ *be given. Then one has for every weak solution* $(X^x, \tilde{W}^x, \mathbb{Q}^x)$ *of equation* (1) *on* $[-r, T]$ *the inequality*

$$\begin{aligned}&\mathbb{E}_{\mathbb{Q}^x} \exp\left(\alpha \sup_{-r\le t\le T} |X^x(t)|^2\right)\\ \le&\frac{C}{\sqrt[4]{1-4\alpha d C_\sigma T}} \exp\left(\frac{\alpha}{1-4\alpha d C_\sigma T}\|x\|_\infty^2 + (16 d C_\sigma T)^{-1}\|x\|_\infty^2\right)\end{aligned}$$

for all $0 \le \alpha < (4dC_\sigma T)^{-1}$ *and a constant* $C = C(d, T, C_\sigma, p, q, \|b\|_{L_p^q(T)}, g_T)$.

Proof. As before, let

$$a^x(t) := \sigma(t, M^x(t))^{-1}\left(B(t, M_t^x) + b(t, M^x(t))\right),\ t \in [0, T].$$

By the assumed conditions, Lemmas 3.3, 3.4, Young's inequality and Hölder's inequality, one has

$$\begin{aligned}&\mathbb{E}_{\mathbb{P}} \exp\left(6\int_0^T |a^x(t)|^2\,\mathrm{d}t\right)\\ \le& C_1\sqrt{\mathbb{E}_{\mathbb{P}}\exp\left(12 C_\sigma \int_0^T |B(t, M_t^x)|\,\mathrm{d}t\right)}\\ \le& C_2\sqrt{\mathbb{E}_{\mathbb{P}}\exp\left((4dC_\sigma T)^{-1}\sup_{-r\le r\le T}|M^x(t)|^2\right)}\\ \le& C_2\exp\left[(8dC_\sigma T)^{-1}\left(\|x\|_\infty^2 - |x(0)|^2\right)\right]\sqrt{\mathbb{E}_{\mathbb{P}}\exp\left(\sup_{0\le t\le T}|M^x(t)|^2\right)}\\ \le& C_3\exp\left((4dC_\sigma T)^{-1}\|x\|_\infty^2\right)\end{aligned}$$

for constants C_1, C_2 and C_3 that only depend on d, T, C_σ, p, q, $\|b\|_{L_p^q(T)}$ and g_T. By Theorem 3.5, one obtains

$$\begin{aligned}&\mathbb{E}_{\mathbb{Q}^x}\left(\alpha \sup_{-r\le t\le T}|X^x(t)|^2\right)\\ =&\mathbb{E}_{\mathbb{P}}\exp\left(\alpha\sup_{-r\le t\le T}|M^x(t)|^2 + \int_0^T a^x(t)^\top \mathrm{d}W(t) - \frac{1}{2}\int_0^T |a^x(t)|^2\,\mathrm{d}t\right)\\ \le&\left[\mathbb{E}_{\mathbb{P}}\exp\left(2\alpha\sup_{-r\le t\le T}|M^x(t)|^2\right)\right]^{\frac{1}{2}}\cdot\left[\mathbb{E}_{\mathbb{P}}\exp\left(2\int_0^T a^x(t)^\top\mathrm{d}W(t) - \int_0^T|a^x(t)|^2\,\mathrm{d}t\right)\right]^{\frac{1}{2}}\\ \le&\left[\mathbb{E}_{\mathbb{P}}\exp\left(2\alpha\sup_{-r\le t\le T}|M^x(t)|^2\right)\right]^{\frac{1}{2}}\cdot\left[\mathbb{E}_{\mathbb{P}}\exp\left(6\int_0^T|a^x(t)|^2\,\mathrm{d}t\right)\right]^{\frac{1}{4}}\\ \le&\frac{C}{\sqrt[4]{1-4\alpha dC_\sigma T}}\exp\left(\frac{\alpha}{1-4\alpha dC_\sigma T}\|x\|_\infty^2 + (16dC_\sigma T)^{-1}\|x\|_\infty^2\right)\end{aligned}$$

for a constant $C = C(d, T, C_\sigma, p, q, \|b\|_{L_p^q(T)}, g_T)$. □

3.3. Stability

The following stability result has essentially been shown in [1], where σ was supposed to be weakly differentiable with

$$\left|\nabla_x \sigma^{i,j}\right| \in L^q_{loc}\left([0,T]; L^p\left(\mathbb{R}^d\right)\right),\ i,j = 1,\dots,d.$$

Formally, condition (C2) does not imply this integrability assumption and for the convenience of the reader, we provide the completely similar proof.

Theorem 3.8. *Assume conditions* (C1), (C2), (C3) *and* (C4)*. Then one has for any* $T_0, R > 0$ *and* $\gamma \geq 1$

$$\mathbb{E}\left\|X_t^x - X_t^y\right\|_\infty^\gamma \leq C\left\|x-y\right\|_\infty^\gamma,\ 0 \leq t \leq T_0$$

for all $x, y \in \mathcal{C}$ *with* $\|x\|_\infty, \|y\|_\infty \leq R$ *and some constant* C *depending only on* γ*,* d*,* p*,* q*,* T_0*,* C_σ*,* $\|b\|_{L^q_p(T_0)}$*,* g_T*,* C_B *and* R*.*

By Theorem A.1, for every $0 < T \leq T_0$, there exists a solution

$$\tilde{u}(\cdot;T) \in \left(H^q_{2,p}(T_0)\right)^d$$

of the coordinatewise PDE system

$$\begin{aligned}\partial_t \tilde{u}(t,x;T) + L_t\tilde{u}(t,x;T) + b(t,x) &= 0,\\ \tilde{u}(T,x;T) &= 0\end{aligned}$$

for all $t \in [0,T]$ and $x \in \mathbb{R}^d$ where

$$L_t v(t,x) := \frac{1}{2}\sum_{i,j,k=1}^d \sigma^{i,k}(t,x)\sigma^{j,k}(t,x)\partial_i\partial_j v(t,x) + b(t,x)\cdot\nabla v(t,x),\ v \in H^q_{2,p}(T_0).$$

Additionally, it holds

$$\sup_{T\in[0,T_0]}\left\|\tilde{u}^i(\cdot;T)\right\|_{H^q_{2,p}(T)} < \infty,\ i = 1,\dots,d$$

and by the embedding Theorem A.2, there exists a uniform δ such that for all $0 \leq S \leq T$ with $T - S \leq \delta$

$$|\tilde{u}(t,x;T) - \tilde{u}(t,y;T)| \leq \frac{1}{2}|x-y|$$

for all $t \in [S,T]$ and $x, y \in \mathbb{R}^d$. Furthermore, the function

$$u(t,x;T) := \tilde{u}(t,x;T) + x$$

satisfies coordinatewise the equation

$$\begin{aligned}\partial_t u(t,x;T) + L_t u(t,x;T) &= 0,\\ u(T,x;T) &= x.\end{aligned}$$

Proof. Choose $\delta > 0$ like above. By induction, it suffices to prove for every $0 \leq S \leq T \leq T_0$ with $T - S \leq \delta$ the implication

$$\begin{aligned} &\mathbb{E}\left\|X_S^x - X_S^y\right\|_\infty^\gamma \leq C_1 \|x-y\|_\infty^\gamma \ \ \forall x, y \in \mathcal{C}, \|x\|_\infty, \|y\|_\infty \leq R \\ \Longrightarrow\ &\mathbb{E}\left\|X_T^x - X_T^y\right\|_\infty^\gamma \leq C_2 \|x-y\|_\infty^\gamma \ \ \forall x, y \in \mathcal{C}, \|x\|_\infty, \|y\|_\infty \leq R \end{aligned}$$

with some constants C_1 and C_2 depending only on γ, d, p, q, T_0, C_σ, $\|b\|_{L_p^q(T_0)}$, g_T, C_B and R. For the sake of simplicity, we write $u(\cdot) := u(\cdot; T)$. Furthermore, define

$$\begin{aligned} Y^x(t) &:= u(t, X^x(t)),\ \ S \leq t \leq T, \\ Y^y(t) &:= u(t, X^y(t)),\ \ S \leq t \leq T. \end{aligned}$$

By the choice of δ, one has for the difference processes $Z(t) := X^x(t) - X^y(t)$ and $\tilde{Z}(t) := Y^x(t) - Y^y(t)$

$$\frac{1}{2}\left|\tilde{Z}(t)\right| \leq |Z(t)| \leq \frac{3}{2}\left|\tilde{Z}(t)\right|,\ \ S \leq t \leq T.$$

Due to Lemma 3.6, Lemma A.3 is applicable, which gives

$$\begin{aligned} \tilde{Z}(t) = &\int_S^t \left(Du(s, X^x(s))B(s, X_s^x) - Du(s, X^y(s))B(s, X_s^y)\right) \mathrm{d}s \\ &+ \int_S^t \left(Du(s, X^x(s))\sigma(s, X^x(s)) - Du(s, X^y(s))\sigma(s, X^y(s))\right) \mathrm{d}W(s) \end{aligned}$$

and consequently

$$\begin{aligned} &\mathrm{d}\left|\tilde{Z}\right|^{2\gamma}(t) \\ =&2\gamma\left|\tilde{Z}(t)\right|^{2\gamma-2}\tilde{Z}(t)^\top\left(Du(t, X^x(t))B(t, X_t^x) - Du(t, X^y(t))B(t, X_t^y)\right)\mathrm{d}t \\ &+ 2\gamma\left|\tilde{Z}(t)\right|^{2\gamma-2}\tilde{Z}(t)^\top\left(Du(t, X^x(t))\sigma(t, X^x(t)) - Du(t, X^y(t))\sigma(t, X^y(t))\right)\mathrm{d}W(t) \\ &+ \gamma\left|\tilde{Z}(t)\right|^{2\gamma-2}\|Du(t, X^x(t))\sigma(t, X^x(t)) - Du(t, X^y(t))\sigma(t, X^y(t))\|_{HS}^2\,\mathrm{d}t \\ &+ 2\gamma(\gamma-1)\left|\tilde{Z}(t)\right|^{2\gamma-4} \\ &\times\left|\left(Du(t, X^x(t))\sigma(t, X^x(t)) - Du(t, X^y(t))\sigma(t, X^y(t))\right)^\top\tilde{Z}(t)\right|^2\mathrm{d}t. \end{aligned}$$

Using the boundedness of Du and condition (C3) gives for $S \leq t_1 \leq t_2 \leq T$

$$\begin{aligned}
&\left|\tilde{Z}(t_2)\right|^{2\gamma} - \left|\tilde{Z}(t_1)\right|^{2\gamma} \\
\leq& c\int_{t_1}^{t_2} \left\|\tilde{Z}_s\right\|_\infty^{2\gamma} \mathrm{d}s \\
+& c\int_{t_1}^{t_2} \left|\tilde{Z}(s)\right|^{2\gamma-1} \|Du(s,X^x(s)) - Du(s,X^y(s))\|_{op} |B(s,X_s^x)| \mathrm{d}s \\
+& c\int_{t_1}^{t_2} \left|\tilde{Z}(s)\right|^{2\gamma-2} \tilde{Z}(s)^\top (Du(s,X^x(s))\sigma(s,X^x(s)) - Du(s,X^y(s))\sigma(s,X^y(s))) \mathrm{d}W(s) \\
+& c\int_{t_1}^{t_2} \left|\tilde{Z}(s)\right|^{2\gamma-2} \|Du(s,X^x(s))\sigma(s,X^x(s)) - Du(s,X^y(s))\sigma(s,X^y(s))\|_{HS}^2 \mathrm{d}s \\
=& I_1 + I_2 + I_3 + I_4
\end{aligned}$$

where $c > 0$ is a constant depending only on γ, d, p, q, T_0, C_σ, $\|b\|_{L_p^q(T_0)}$, g_T, C_B and R. The idea is to apply the stochastic Gronwall Lemma A.5. To get rid of the badly behaving terms I_2 and I_4, one can use a suitable multiplier of the form $e^{-A(t)}$ - as in [8] - where A is an adapted, continuous process. Here, we choose

$$\begin{aligned}
A(t) := c\int_S^t |B(s,X_s^x)| \frac{\|Du(s,X^x(s)) - Du(s,X^y(s))\|_{op}}{\left|\tilde{Z}(s)\right|} \mathbb{1}_{\tilde{Z}(s)\neq 0} \mathrm{d}s \\
+ c\int_S^t \frac{\|Du(s,X^x(s))\sigma(s,X^x(s)) - Du(s,X^y(s))\sigma(s,X^y(s))\|_{HS}^2}{\left|\tilde{Z}(s)\right|^2} \mathbb{1}_{\tilde{Z}(s)\neq 0} \mathrm{d}s
\end{aligned}$$

for $S \leq t \leq T$. To show that A is indeed well defined, it suffices to show the existence of a constant $\hat{C} = \hat{C}(\gamma, d, p, q, C_\sigma, T_0, \|b\|_{L_p^q(T_0)}, g_T, C_B, R) \geq 0$ such that

$$\mathbb{E} \exp\left(\frac{1}{2}A(T)\right) \leq \hat{C}.$$

Since u belongs coordinatewise to $H_{2,p}^q(T_0)$ and by condition (C2), it holds

$$(Du \cdot \sigma)^{i,j} \in L^q\left(T_0; W^{1,p}\left(\mathbb{R}^d\right)\right), \ i,j = 1,\ldots,d.$$

Additionally, $C_c^\infty\left(\mathbb{R}^{d+1}\right)$ is dense in $L^q\left(T_0; W^{1,p}\left(\mathbb{R}^d\right)\right)$. Hence, by Young's inequality, Lemma 3.6 and Lemma 3.7, it suffices to show for all $\tilde{R} > 0$ the existence of a constant $C_{\tilde{R}} = C_{\tilde{R}}(d, p, q, C_\sigma, T_0, \|b\|_{L_p^q(T_0)}, g_T, C_B, R)$ such that

$$\mathbb{E} \exp\left(\int_S^T \frac{|f(s,X^x(s)) - f(s,X^y(s))|^2}{\left|\tilde{Z}(s)\right|^2} \mathbb{1}_{\tilde{Z}(s)\neq 0} \mathrm{d}s\right) \leq C_{\tilde{R}}$$

for all $f \in C^\infty\left(\mathbb{R}^{d+1}\right)$ with $\|f\|_{L^q\left(T_0;W^{1,p}\left(\mathbb{R}^d\right)\right)} \leq \tilde{R}$. By Lemmas 3.6 and A.4, one obtains

$$\begin{aligned}
&\mathbb{E}\exp\left(\int_S^T \frac{|f(s,X^x(s)) - f(s,X^y(s))|^2}{\left|\tilde{Z}(s)\right|^2} \mathbb{1}_{\tilde{Z}(s)\neq 0}\mathrm{d}s\right) \\
\leq&\mathbb{E}\exp\left(C_d^2\int_S^T \left(\mathcal{M}\left|\nabla f\right|(X^x(s)) + \mathcal{M}\left|\nabla f\right|(X^y(s))\right)^2\mathrm{d}s\right) \\
\leq& C_{\tilde{R}}
\end{aligned}$$

where $C_{\tilde{R}} = C_{\tilde{R}}(d,p,q,C_\sigma,T_0,\|b\|_{L^q_p(T_0)},g_T,C_B,R)$. By the Itō formula, it holds

$$e^{-A(t)}\left|\tilde{Z}(t)\right|^{2\gamma} \leq \left|\tilde{Z}(S)\right|^{2\gamma} + c\int_S^t e^{-A(s)}\left\|\tilde{Z}_s\right\|_\infty^{2\gamma}\mathrm{d}s + \text{local martingale}.$$

Applying the stochastic Gronwall Lemma A.5 gives

$$\mathbb{E}\left[\sup_{S\leq t\leq T} e^{-\frac{1}{2}A(t)}\left|\tilde{Z}(t)\right|^\gamma\right] \leq \tilde{C}\mathbb{E}\left\|\tilde{Z}_S\right\|_\infty^\gamma \leq \tilde{C}C_1\|x-y\|_\infty^\gamma$$

for a constant $\tilde{C} = \tilde{C}(\gamma,d,p,q,C_\sigma,T_0,\|b\|_{L^q_p(T_0)},g_T,C_B,R)$. Due to the estimates from above, the Cauchy-Schwarz inequality and by redefining $\gamma := 2\gamma$, one finally obtains

$$\begin{aligned}
&\mathbb{E}\left[\sup_{S\leq t\leq T}|Z(t)|^\gamma\right] \\
\leq&\left(\mathbb{E}e^{\frac{1}{2}A(T)}\right)^{\frac{1}{2}}\left[\mathbb{E}\left(\sup_{S\leq t\leq T} e^{-\frac{1}{2}A(t)}|Z(t)|^{2\gamma}\right)\right]^{\frac{1}{2}} \\
\leq& C_2\|x-y\|_\infty^\gamma
\end{aligned}$$

for some constant $C_2 = C_2(\gamma,d,p,q,C_\sigma,T_0,\|b\|_{L^q_p(T_0)},g_T,C_B,R)$. □

3.4. Strong Feller Property

The following theorem is a consequence of a log-Harnack inequality that has been shown in [24] and requires the Lipschitz-continuity of σ in space.

Theorem 3.9. *Assume condition* (C2). *Then one has for all $t > r$*

$$\lim_{y\to x}\mathbb{E}f(M_t^y) = \mathbb{E}f(M_t^x)\ \forall f\in B_b(\mathcal{C}).$$

Although, the drift-free equation has no delay, we need the strong Feller property with respect to the state space of path segments $\mathcal{C}$. For Theorem 3.9, we do not know so far whether one can weaken the Lipschitz condition on σ to the “usual” assumption

$$\left|\nabla_x\sigma^{i,j}\right| \in L^q_{loc}\left([0,T];L^p\left(\mathbb{R}^d\right)\right),\ i,j=1,\ldots,d.$$

Then the following results would still hold even with that weaker assumption.

Lemma 3.10. *Assume condition* (C2). *Then one has*

$$\lim_{y\to x} \mathbb{E} \int_0^T |b(t, M^x(t)) - b(t, M^y(t))|^2 \,\mathrm{d}t = 0.$$

Proof. By Theorems 1.7 and 3.9, one has for all $t > 0$

$$\lim_{y\to x} \mathbb{E}\, |f(M^x(t)) - f(M^y(t))| = 0 \;\forall f \in B_b(\mathbb{R}^d).$$

Therefore, one has for all $f \in B_b\left([0,T] \times \mathbb{R}^d\right)$

$$\lim_{y\to x} \mathbb{E} \int_0^T |f(t, M^x(t)) - f(t, M^y(t))| \,\mathrm{d}t = 0.$$

Consequently, $b(\cdot, M^{x_n}(\cdot))$ converges to $b(\cdot, M^x(\cdot))$ in measure with respect to $\mathbb{P} \otimes \lambda_{|[0,T]}$ for each sequence $(x_n)_{n\in\mathbb{N}} \subset \mathcal{C}$ converging to x. By Lemma 3.2, it follows

$$\lim_{\alpha\to\infty} \sup_{y\in\mathcal{C}} \mathbb{E} \int_0^T \mathbb{1}_{|b(t,M^y(t))|\geq\alpha} \, |b(t, M^y(t))|^2 \,\mathrm{d}t = 0.$$

Hence, $\left\{|b(t, M^y(t))|^2 : y \in \mathcal{C}\right\}$ is uniformly integrable and the stated L^2-convergence follows. ☐

Proof of Theorem 1.6. Let $t > r$. Due to Theorems 1.7 and 3.8, it suffices to show that

$$\lim_{y\to x} \mathbb{E}_{\mathbb{Q}^y} f(X_t^y) = \mathbb{E}_{\mathbb{Q}^x} f(X_t^x) \;\forall f \in B_b(\mathcal{C}).$$

Let $f \in B_b(\mathcal{C})$, then one has by Theorem 3.5

$$\begin{aligned}
&\mathbb{E}_{\mathbb{Q}^x} f(X_t^x) - \mathbb{E}_{\mathbb{Q}^y} f(X_t^y) \\
=&\mathbb{E}_{\mathbb{P}}(D^x(t) f(M_t^x)) - \mathbb{E}_{\mathbb{P}}(D^y(t) f(M_t^y)) \\
=&\mathbb{E}_{\mathbb{P}}[D^x(t)(f(M_t^x) - f(M_t^y))] + \mathbb{E}_{\mathbb{P}}[(D^x(t) - D^y(t)) f(M_t^y)] \\
\leq&\mathbb{E}_{\mathbb{P}}[D^x(t)(f(M_t^x) - f(M_t^y))] + \|f\|_\infty \mathbb{E}_{\mathbb{P}}\, |D^x(t) - D^y(t)|
\end{aligned}$$

where we define for every $z \in \mathcal{C}$

$$\begin{aligned}
a^z(t) &:= \sigma(t, M^z(t))^{-1} \left(B(t, M_t^z) + b(t, M^z(t))\right), \\
D^z(t) &:= \exp\left(\int_0^t a^z(s)^\top \mathrm{d}W(s) - \frac{1}{2}\int_0^t |a^z(s)|^2 \,\mathrm{d}s\right).
\end{aligned}$$

By condition (C2), Itō's formula and the stochastic Gronwall Lemma A.5, it holds

$$\lim_{n\to\infty} M_t^{x_n} = M_t^x \text{ in probability}$$

for each sequence $(x_n)_{n\in\mathbb{N}} \subset \mathcal{C}$ converging to x. Applying Theorems 1.7 and 3.9, gives

$$\lim_{y\to x} \mathbb{E}_{\mathbb{P}}\, |f(M_t^y) - f(M_t^x)| = 0$$

and in particular,

$$\lim_{n\to\infty} D^x(t) f(M_t^{x_n}) = D^x(t) f(M_t^x) \text{ in probability}$$

for each sequence $(x_n)_{n\in\mathbb{N}} \subset \mathcal{C}$ converging to x. By the dominated convergence theorem, it follows

$$\lim_{y\to x} \mathbb{E}_{\mathbb{P}}[D^x(t)(f(M_t^y) - f(M_t^x))] = 0.$$

Consequently, it remains to show that

$$\lim_{y\to x} \mathbb{E}_{\mathbb{P}} \left|D^y(t) - D^x(t)\right| = 0.$$

Since one has $\mathbb{E}_{\mathbb{P}} D^z(t) = 1$ for all $z \in \mathcal{C}$, it suffices to show for each sequence $(x_n)_{n\in\mathbb{N}} \subset \mathcal{C}$ converging to x

$$\lim_{n\to\infty} D^{x_n}(t) = D^x(t) \text{ in probability.}$$

This can be seen as follows: then by Fatou's lemma,

$$2 - \lim_{n\to\infty} \mathbb{E}_{\mathbb{P}} \left|D^{x_n}(t) - D^x(t)\right| = \lim_{n\to\infty} \mathbb{E}_{\mathbb{P}} \left(D^x(t) + D^{x_n}(t) - \left|D^{x_n}(t) - D^x(t)\right|\right) \geq 2$$

holds and the L^1-convergence would be an immediate consequence. Therefore, it is sufficient to show

$$\lim_{y\to x} \mathbb{E}_{\mathbb{P}} \int_0^t \left|a^y(s) - a^x(s)\right|^2 \mathrm{d}s = 0$$

by the martingale isometry. One has

$$\begin{aligned}
&\mathbb{E}_{\mathbb{P}} \int_0^t \left|a^y(s) - a^x(s)\right|^2 \mathrm{d}s \\
\leq &2\mathbb{E}_{\mathbb{P}} \int_0^t \left\|\sigma\left(s, M^y(s)\right)^{-1} - \sigma\left(s, M^x(s)\right)^{-1}\right\|_{op}^2 \left|B(s, M_s^x) + b(t, M^x(s))\right|^2 \mathrm{d}s \\
&+ 2C_\sigma \mathbb{E}_{\mathbb{P}} \int_0^t \left|B(s, M_s^y) + b(s, M^y(s)) - B(s, M_s^x) - b(s, M^x(s))\right|^2 \mathrm{d}t.
\end{aligned}$$

The second term converges to zero by condition (C3), Theorem 1.7, Lemma 3.4, Theorem 3.9 and Lemma 3.10. Moreover, for each sequence $(x_n)_{n\in\mathbb{N}} \subset \mathcal{C}$ converging to x,

$$\lim_{n\to\infty} \left\|\sigma\left(\cdot, M^{x_n}(\cdot)\right)^{-1} - \sigma\left(\cdot, M^x(\cdot)\right)^{-1}\right\|_{op} = 0 \text{ in measure w.r.t. } \mathbb{P} \otimes \lambda_{|[0,t]}$$

holds by Theorem 3.8, the continuity of σ in space and the continuity of the inverting map $A \mapsto A^{-1}$ on the space of invertible matrices. Additionally, one can bound the first integrand by

$$2C_\sigma \left|B(\cdot, M_\cdot^x) + b(\cdot, M^x(\cdot))\right|^2,$$

which is $\mathbb{P} \otimes \lambda_{|[0,t]}$-integrable by Lemma 3.2. Consequently, one can apply the dominated convergence theorem and the proof is complete. □

A. Appendix

Theorem A.1. *Assume conditions* (C1) *and* (C2). *Then for any* $T > 0$ *and* $f \in L_p^q(T)$, *there exists a unique solution* $u \in H_{2,p}^q(T)$ *of the following PDE*

$$\begin{aligned} \partial_t u(t,x) + L_t u(t,x) + f(t,x) &= 0, \\ u(T,x) &= 0 \end{aligned}$$

with the bound

$$\|u\|_{H_{2,p}^q(S,T)} \leq C \|f\|_{L_p^q(S,T)}$$

for any $S \in [0,T]$ *and some constant* $C = C(T, C_\sigma, p, q, \|b\|_{L_p^q(T)}) > 0$.

Proof. See [26]. □

Theorem A.2. *Let* $p, q \in (1,\infty)$, $T > 0$ *and* $u \in H_{2,p}^q(T)$.

1. *If* $\frac{d}{p} + \frac{2}{q} < 2$, *then* u *is a bounded Hölder continuous function on* $[0,T] \times \mathbb{R}^d$ *and for any* $0 < \varepsilon,\ \delta \leq 1$ *satisfying*

$$\varepsilon + \frac{d}{p} + \frac{2}{q} < 2, \quad 2\delta + \frac{d}{p} + \frac{2}{q} < 2,$$

there exists a constant $N = N(p,q,\varepsilon,\delta)$ *such that*

$$\begin{aligned} |u(t,x) - u(s,x)| &\leq N |t-s|^\delta \|u\|_{\mathbb{H}_{2,p}^q(T)}^{1-\frac{1}{q}-\delta} \|\partial_t u\|_{L_p^q(T)}^{\frac{1}{q}+\delta}, \\ |u(t,x)| + \frac{|u(t,x)-u(t,y)|}{|x-y|^\varepsilon} &\leq N T^{-\frac{1}{q}} \left(\|u\|_{\mathbb{H}_{2,p}^q(T)} + T \|\partial_t u\|_{L_p^q(T)} \right) \end{aligned}$$

for all $s,t \in [0,T]$ *and* $x,y \in \mathbb{R}^d, x \neq y$.

2. *If* $\frac{d}{p} + \frac{2}{q} < 1$, *then* ∇u *is a bounded Hölder continuous function on* $[0,T] \times \mathbb{R}^d$ *and for any* $\varepsilon \in (0,1)$ *satisfying*

$$\varepsilon + \frac{d}{p} + \frac{2}{q} < 1,$$

there exists a constant $N = N(p,q,\varepsilon)$ *such that*

$$\begin{aligned} |\nabla u(t,x) - \nabla u(s,x)| &\leq N |t-s|^\delta \|u\|_{\mathbb{H}_{2,p}^q(T)}^{1-\frac{1}{q}-\frac{\varepsilon}{2}} \|\partial_t u\|_{L_p^q(T)}^{\frac{1}{q}+\frac{\varepsilon}{2}}, \\ |\nabla u(t,x)| + \frac{|\nabla u(t,x)-\nabla u(t,y)|}{|x-y|^\varepsilon} &\leq N T^{-\frac{1}{q}} \left(\|u\|_{\mathbb{H}_{2,p}^q(T)} + T \|\partial_t u\|_{L_p^q(T)} \right) \end{aligned}$$

for all $s,t \in [0,T]$ *and* $x,y \in \mathbb{R}^d, x \neq y$.

Proof. See [7, p. 22, 23, 36]. □

In the next lemma we identify every $u \in H_{2,p}^q$ with its regular version.

Lemma A.3 (Itō formula for $H^q_{2,p}$-functions)**.** *Let $T > 0$, $p > 1$ and $q > 1$ satisfying (2). Let $X : \Omega \times [0,T] \to \mathbb{R}^d$ be a semimartingale on some filtrated probability space $\left(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0}\right)$ of the form*

$$\mathrm{d}X(t) = b(t)\mathrm{d}t + \sigma(t)\mathrm{d}W(t)$$

where W is a d-dimensional Brownian motion, $b : \Omega \times [0,T] \to \mathbb{R}^d$ and $\sigma : \Omega \times [0,T] \to \mathbb{R}^{d\times d}$ are progressively measurable with

$$\mathbb{P}\left(\|b\|_{L^1[0,T]} + \left\|a^{i,j}\right\|_{L^\delta[0,T]} < \infty\right) = 1, \ i,j = 1,\dots,d$$

for some $1 < \delta \leq \infty$ where $a := \sigma\sigma^\top$. Furthermore, assume that there exists a constant $C > 0$ with

$$\mathbb{E}\int_0^T f(t, X(t))\mathrm{d}t \leq C\,\|f\|_{L^{q/\delta^*}_{p/\delta^*}(T)}$$

for all $f \in L^{q/\delta^*}_{p/\delta^*}(T)$ where δ^* denotes the conjugate exponent of δ. Then for any $u \in H^q_{2,p}(T)$, the Itō formula holds, i.e.

$$\begin{aligned} u(t, X(t)) - u(0, X(0)) = &\int_0^t \partial_t u(s, X(s))\mathrm{d}s + \int_0^t \nabla u(s, X(s))^\top b(s)\mathrm{d}s \\ &+ \int_0^t \nabla u(s, X(s))^\top \sigma(s)\mathrm{d}W(s) \\ &+ \frac{1}{2}\sum_{i,j=1}^d \int_0^t \partial_i\partial_j u(s, X(s)) a^{i,j}(s)\mathrm{d}s. \end{aligned}$$

Proof. See [1]. □

Let ϕ be a locally integrable function on $\mathbb{R}^d$. The Hardy-Littlewood maximal function is defined by

$$\mathcal{M}\phi(x) := \sup_{0<r<\infty} \frac{1}{|B_r|}\int_{B_r} \phi(x+y)\mathrm{d}y$$

where B_r is the Euclidean ball of radius r. The following result is cited from [26].

Lemma A.4.

1. *There exists a constant $C_d > 0$ such that for all $\phi \in C^\infty\left(\mathbb{R}^d\right)$ and $x, y \in \mathbb{R}^d$,*

$$|\phi(x) - \phi(y)| \leq C_d\,|x-y|\left(\mathcal{M}\,|\nabla\phi|\,(x) + \mathcal{M}\,|\nabla\phi|\,(y)\right).$$

2. *For any $p > 1$, there exists a constant $C_{d,p}$ such that for all $\phi \in L^p\left(\mathbb{R}^d\right)$,*

$$\|\mathcal{M}\phi\|_{L^p} \leq C_{d,p}\,\|\phi\|_{L^p}.$$

Lemma A.5. *Let Z be an adapted non-negative stochastic process with continuous paths defined on $[0,\infty)$ that satisfies the inequality*

$$Z(t) \leq K \int_0^t \sup_{0\leq r\leq s} Z(r)\mathrm{d}s + M(t) + C,$$

where $C \geq 0$, $K \geq 0$ and M is a continuous local martingale with $M(0) = 0$. Then for each $0 < p < 1$, there exist universal finite constants $c_1(p)$, $c_2(p)$ (not depending on K, C, T and M) such that

$$\mathbb{E}\left[\sup_{0\leq t\leq T} Z(t)^p\right] \leq C^p c_2(p) e^{c_1(p)KT} \text{ for every } T \geq 0.$$

Proof. See [23]. □

B. The Strict Topology on $\mathrm{C_b(E)}$

In this subsection we want to consider the strict topology for the function space $C_b(E)$ of bounded, continuous functions on a polish space E as a nontrivial example where the results given before are applicable.

In this subsection, we assume that E is a polish space. Equipping $C_b(E)$ with the usual supremum norm might have some drawbacks if E is not (locally) compact. A well-known result is the following

Proposition B.1. $(C_b(E), \|\cdot\|_\infty)$ *is separable iff E is compact.*

Additionally, if E is not locally compact, the dual space of $(C_b(X), \|\cdot\|)$ may not coincide with the space of complex Borel measures on E. That might give rise to consider different topologies on $C_b(E)$. The strict topology on $C_b(E)$ is defined as follows.

Definition B.2. Define

$$H^+(E) := \{u : E \to \mathbb{R}_{\geq 0} : \{u \geq \alpha\} \text{ compact for all } \alpha > 0\}.$$

Then the strict topology β on $C_b(E)$ shall be generated by the seminorms

$$\rho_u(f) := \|uf\|_\infty, \ f \in C_b(E).$$

Remark B.3. It turns out that a sequence converges with respect to the strict topology iff the sequence converges uniformly on compact sets and is uniformly bounded (both with respect to the supremum norm).

The strict topology has a rich structure and is discussed deeply in the context of Markov processes and Feller semigroups in [20]. A remarkable property is the following: the topology β is the finest locally convex topology such that the dual space coincides with the space of complex Borel measures on E.

Definition B.4. A collection $\mathcal{N}$ of subsets of E is called a network if for any $x \in E$ and open $O \subseteq E$ with $x \in O$ there exists some $N \in \mathcal{N}$ with $x \in N \subseteq O$.

Lemma B.5. *The space $(C(E), \beta)$ has a countable network. In particular, it is hereditarily Lindelöf.*

Proof. The approach is analogous to the one for the topology of pointwise convergence, which can be found in Lemma 7.1 in [19]. Let $\mathfrak{B}$ be a countable base for E. Then define for $B \in \mathfrak{B}$, $m \in \mathbb{N}$, $a, b \in \mathbb{Q}$ with $a < b$

$$[B, a, b, m] := \{h \in C_b(E) : h(B) \subseteq (a, b), \|h\|_\infty < m\}$$

and

$$\tilde{\mathcal{N}} := \{[B, a, b, m] : B \in \mathfrak{B}, a, b \in \mathbb{Q}, a < b, m \in \mathbb{N}\}.$$

Now,

$$\mathcal{N} := \left\{\bigcap_{i=1}^{n} A_i : A_1, \ldots, A_n \in \tilde{\mathcal{N}}, n \in \mathbb{N}\right\}$$

is countable and the candidate for the claimed network. Let $f \in C_b(E)$, $u \in H^+(E)$ and $\varepsilon > 0$ be given. Then one has to show that there exists some $N \in \mathcal{N}$ such that

$$f \in N \subseteq \{g \in C_b(E) : \|u(g - f)\|_\infty < \varepsilon\}.$$

The first step is to use the compactness property of u: define

$$m := \inf\{n \in \mathbb{N} : n \geq \|f\|_\infty\}$$

and

$$K := \left\{x \in E : |u(x)| \geq \frac{\varepsilon}{2m}\right\}.$$

By construction, it holds for each $B \in \mathfrak{B}$, $a, b \in \mathbb{Q}$ with $a < b$

$$|u(x)(h(x) - f(x))| < \varepsilon \ \forall x \in E \setminus K, h \in [B, (a, b), m].$$

If $K = \emptyset$, the proof would be finished. Thus, assume $K \neq \emptyset$ and define

$$\tilde{\varepsilon} := \frac{\varepsilon}{\|u\|_\infty}.$$

So, it remains to show that there exists an $N \in \mathcal{N}$ with $f \in N$ and

$$|h(x) - f(x)| < \tilde{\varepsilon} \ \forall x \in K, h \in N.$$

Since f is continuous, there exists for every $x \in K$ a $B_x \in \mathfrak{B}$ with

$$f(B_x) \subseteq (f(x) - \tilde{\varepsilon}/4, f(x) + \tilde{\varepsilon}/4).$$

Additionally, K is compact. Consequently, there exist $x_1, \dots, x_n$, $n \in \mathbb{N}$ with

$$K \subseteq \bigcup_{i=1}^{n} B_{x_i}.$$

Now, choose suitable $a_1, b_1, \dots, a_n, b_n$ with

$$f(x_i) < f(x_i) - \tilde{\varepsilon}/2 < a_i < f(x_i) - \tilde{\varepsilon}/4 < f(x_i) + \tilde{\varepsilon}/4 < b_i < f(x_i) + \tilde{\varepsilon}/2, \; i = 1 \dots n.$$

Then it holds $f \in \bigcap_{i=1}^{n} [B_{x_i}, a_i, b_i, m] \in \mathcal{N}$ and

$$|h(x) - f(x)| < \tilde{\varepsilon} \;\forall x \in K, h \in \bigcap_{i=1}^{n} [B_{x_i}, a_i, b_i, m].$$

$\square$

In the following corollary, the space $C_b(E)$ will be implicitly equipped with the strict topology β.

Corollary B.6. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be some probability space and $X, X_n : \Omega \to C_b(E)$, $n \in \mathbb{N}$ be measurable maps. Then the following statements are equivalent*

1. *a)* $\lim_{n\to\infty} \mathbb{P}\left(\rho_u(X - X_n) > \varepsilon\right) = 0 \;\forall \varepsilon > 0, u \in H^+(E)$,

   b) $\lim_{n\to\infty} \mathbb{P}_{X_n}(O) = \mathbb{P}_X(O)$, *for all open* $O \subset C_b(E)$.
2. $\lim_{n\to\infty} \mathbb{E}\left|f(X) - f(X_n)\right| = 0 \;\forall f \in B_b(E)$.

Remark B.7. The strict topology is not metrizable if E is not compact. This can be seen as follows: assume that $(C(E), \beta)$ is metrizable and E is not compact. Then the zero function has a countable neighborhood base and there is a sequence $(x_n)_{n\in\mathbb{N}}$ in E that has no cluster points with $x_n \neq x_m$ if $n \neq m$. Thus, there exists a countable set $(u_i)_{i\in\mathbb{N}} \subset H^+(E)$ such that

$$\forall u \in H^+(E) \;\exists i \in \mathbb{N} : u(x_n) \leq u_i(x_n) \;\forall n \in \mathbb{N}.$$

Additionally, one has for every $u \in H^+(E)$

$$\lim_{n\to\infty} u(x_n) = 0.$$

Now, one can construct a $\hat{u} \in H^+(E)$ that contradicts to the inequality above. Define inductively

$$n_1 := 1, \; n_{i+1} := \inf\left\{n \in \mathbb{N} : n > n_i, u_{i+1}(x_n) \leq 1/(i+1)\right\}$$

and

$$\begin{aligned} \hat{u}\left(x_{n_i}\right) &:= u_i\left(x_{n_i}\right) + 1/i, \\ \hat{u}(y) &:= 0, \; y \in E \text{ with } \nexists\, i \in \mathbb{N} : \; y = x_{n_i}. \end{aligned}$$

Indeed, the function $\hat{u}$ is well defined and it holds $\hat{u} \in H^+(E)$ with

$$\hat{u}\left(x_{n_i}\right) > u_i\left(x_{n_i}\right), \; i \in \mathbb{N}.$$

References

[1] S. Bachmann. Well-posedness and stability for a class of stochastic delay differential equations with singular drift. *Stochastics and Dynamics*, 0(0):1850019, 0.

[2] S. Blei and H.-J. Engelbert. One-dimensional stochastic differential equations with generalized and singular drift. *Stochastic Processes and their Applications*, 123(12):4337 – 4372, 2013.

[3] F. Cordoni and L. Di Persio. Stochastic reaction-diffusion equations on networks with dynamic time-delayed boundary conditions. *Journal of Mathematical Analysis and Applications*, 451(1):583 – 603, 2017.

[4] F. Cordoni, L. Di Persio, and I. Oliva. A nonlinear kolmogorov equation for stochastic functional delay differential equations with jumps. *Nonlinear Differential Equations and Applications NoDEA*, 24(2):16, Mar 2017.

[5] R. M. Dudley. *Real Analysis and Probability*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2 edition, 2002.

[6] A. Es-Sarhir, M.-K. von Renesse, and M. Scheutzow. Harnack inequality for functional sdes with bounded memory. *Electron. Commun. Probab.*, 14:560–565, 2009.

[7] E. Fedrizzi. Uniqueness and flow theorems for solutions of SDEs with low regularity of the drift. *tesi di Laurea in Matematica, Università di Pisa*, 2009.

[8] E. Fedrizzi and F. Flandoli. Pathwise uniqueness and continuous dependence of SDEs with non-regular drift. *Stochastics*, 83(3):241–257, 2011.

[9] L. Gawarecki and V. Mandrekar. *Stochastic Differential Equations with Discontinuous Drift*, pages 185–200. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.

[10] I. Gyöngy and T. Martínez. On stochastic differential equations with locally unbounded drift. *Czechoslovak Math. J.*, 51(126)(4):763–783, 2001.

[11] X. Huang. Strong solutions for functional sdes with singular drift. *Stochastics and Dynamics*, 0(0):1850015, 0.

[12] N. V. Krylov and M. Röckner. Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Related Fields*, 131(2):154–196, 2005.

[13] R. S. Liptser and A. N. Shiryaev. *Statistics of Random Processes: I. General Theory*, pages 286–297. Springer, 2001.

[14] B. Maslowski and J. Seidler. Probabilistic approach to the strong feller property. *Probability Theory and Related Fields*, 118(2):187–210, Oct 2000.

[15] N. I. Portenko. *Generalized diffusion processes*, volume 83 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1990. Translated from the Russian by H. H. McFaden.

[16] G. Da Prato, K.D. Elworthy, and J. Zabczyk. Strong feller property for stochastic semilinear equations. *Stochastic Analysis and Applications*, 13(1):35–45, 1995.

[17] M. Rutkowski. Stochastic differential equations with singular drift. *Statistics & Probability Letters*, 10(3):225 – 229, 1990.

[18] M. Scheutzow. A stochastic gronwall lemma. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 16(02):1350019, 2013.

[19] D. Shakhmatov and J. Spěvák. Group-valued continuous functions with the topology of pointwise convergence. *Topology and its Applications*, 157(8):1518 – 1540, 2010. Advances in Set-Theoretic Topology: Proceedings of the Conference in Honour of Professor Tsugunori Nogura on his Sixtieth Birthday (9–19 June 2008, Erice, Sicily, Italy).

[20] J. A. van Casteren. *Markov Processes, Feller Semigroups and Evolution Equations.* Series on concrete and applicable mathematics. World Scientific, 2011.

[21] A. W. van der Vaart and J. Wellner. *Weak Convergence and Empirical Processes: With Applications to Statistics.* Springer Series in Statistics. Springer, 1996.

[22] A. Y. Veretennikov. Strong solutions of stochastic differential equations. *Teor. Veroyatnost. i Primenen.*, 24(2):348–360, 1979.

[23] M.-K. von Renesse and M. Scheutzow. Existence and uniqueness of solutions of stochastic functional differential equations. *Random Oper. Stoch. Equ.*, 18(3):267–284, 2010.

[24] F.-Y. Wang and C. Yuan. Harnack inequalities for functional SDEs with multiplicative noise and applications. *Stochastic Process. Appl.*, 121(11):2692–2710, 2011.

[25] T. Yamada and S. Watanabe. On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.*, 11(1):155–167, 1971.

[26] X. Zhang. Stochastic homeomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients. *Electron. J. Probab.*, 16:no. 38, 1096–1116, 2011.

[27] A. K. Zvonkin. A transformation of the phase space of a diffusion process that will remove the drift. *Mat. Sb. (N.S.)*, 93(135):129–149, 152, 1974.

On the Strong Feller Property and Well-Posedness for SDEs with Functional, Locally Unbounded Drift

Stefan Bachmann

Email: bachmann@math.uni-leipzig.de

Institut für Mathematik, Universität Leipzig,
Augustusplatz 10, 04109 Leipzig, Germany

December 8, 2019

Abstract. We study functional stochastic differential equations with a locally unbounded, functional drift focusing on well-posedness, stability and the strong Feller property. Following the non-functional case, we consider integrability conditions and only need minimal continuity assumptions. Our approach is mainly based on Zvonkin's transformation [18] and the convergence concept for random variables in topological spaces in [2], which extends the probabilistic approach of Maslowski and Seidler [12]. Our arguments for the strong Feller property are mostly probabilistic, relatively elementary and can still deal with non-regular drifts. This allows extensions in various ways and are applicable in different, more complex situations.

Keywords: stochastic delay differential equations, stochastic functional differential equations, retarded differential equations, strong Feller property, pathwise uniqueness, regularization by white noise, singular drift, unbounded drift

MSC 2010: primary 34K50; secondary 60B10, 60B12, 60H10.

1. Introduction

In this paper, we consider stochastic functional differential equations of the following form

$$\begin{aligned} \mathrm{d}X^x(t) &= B(t, X^x)\mathrm{d}t + \sigma(t, X^x(t))\mathrm{d}W(t) \\ X_0 &= x \in C\left([-r, 0], \mathbb{R}^d\right) \end{aligned} \tag{1}$$

where W is a d-dimensional Brownian motion, $B : \mathbb{R}_{\geq 0} \times C(\mathbb{R}_{\geq -r}, \mathbb{R}^d) \to \mathbb{R}^d$ is non-anticipating and $\sigma : \mathbb{R}_{\geq 0} \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ is measurable, bounded, non-degenerate and Lipschitz in space.

Non-functional stochastic differential equations (SDEs) with discontinuous drift have been extensively studied: Portenko [13], Veretennikov [15] and Zvonkin [18] considered - among other things - well-posedness for SDEs with bounded, discontinuous drift terms. Krylov and Röckner have shown existence and uniqueness for locally unbounded drifts and constant, non-degenerate diffusion coefficients in [10]. Singular SDEs with non-constant, non-degenerate diffusion matrices have been studied by Martínez, Gyōngy [7] and Zhang [17]. Additionally, there are numerous results for the strong Feller property for non-functional, singular SDEs with the Euclidean state space $\mathbb{R}^d$ (i.e. [17]).

However, we are interested in the strong Feller property for functional SDEs with the state space of path segments $C\left([-r,0], \mathbb{R}^d\right)$ for some $r > 0$. Es-Sarhir, von Renesse and Scheutzow established a Harnack-inequality under Lipschitz conditions and constant, non-degenerate diffusion matrices in [4], which implies the strong Feller property. Wang and Yuan proved a log-Harnack inequality for non-constant, non-degenerate diffusion coefficients in [16]. In [1] and [8] well-posedness has been considered for SDEs with a drift consisting of a functional part and a non-functional, locally unbounded part. The strong Feller property has been shown in [2].

To prove the strong Feller property for functional, locally unbounded drifts, we use the following convergence concept for random variables which has been invented by the author in [2].

Theorem 1.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be some probability space and (E, d) be a metric space. Furthermore, let $X, X_n : \Omega \to E$, $n \in \mathbb{N}$ be measurable maps. Then the statement*

1. *a)* $\lim_{n\to\infty} \mathbb{P}^*\left(d\left(X, X_n\right) \geq \varepsilon\right) = 0 \ \forall \varepsilon > 0,$

   b) $\lim_{n\to\infty} \mathbb{P}_{X_n}(O) = \mathbb{P}_X(O)$ *for all open* $O \subset E$

implies

2. $\lim_{n\to\infty} \mathbb{E}\left|f(X) - f(X_n)\right| = 0$ *for all bounded, measurable* $f : E \to \mathbb{R}$

where $\mathbb{P}^*$ denotes the outer measure of $\mathbb{P}$. Additionally, if there exists some null set $N \subset \Omega$ such that $X(\Omega \setminus N)$ is separable, then the converse implication is also true.

Proof. See Theorem 1.7 in [2]. □

Although it seems to be a general probabilistic but rather abstract result, it can be applied to the framework of stochastic differential equations. It allows us to show convergence of the Girsanov densities

$$\lim_{y\to x} \frac{\mathrm{d}X^y_{|[-r,T]}}{\mathrm{d}M^y_{|[-r,T]}} = \frac{\mathrm{d}X^x_{|[-r,T]}}{\mathrm{d}M^x_{|[-r,T]}} \text{ in probability}$$

for $T > r$ where M^x denotes the solution of equation (1) without drift, i.e.

$$\begin{aligned} \mathrm{d}M^x(t) &= \sigma(t, M^x(t))\mathrm{d}W(t) \\ M_0 &= x \end{aligned}$$

Roughly speaking, this is the main step for proving the strong Feller property. For that, Theorem 1.1 is extremely useful, especially to deal with discontinuous coefficients, which extends the probabilistic approach of Maslowski and Seidler [12]. Remarkably, the coefficients in the drift-free equation do not depend on the past. For proving well-posedness and stability we combine Zvonkin's transformation [18] and combine it with the convergence result Theorem 1.1. In both cases, we extensively make use of Krylov's estimate for semimartingales [9].

Notation 1.2. If not stated otherwise, W will be a d-dimensional Brownian motion on some arbitrary but fixed probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and every strong solution shall be defined on this space.

However, weak solutions of equation (1) might be defined on different filtrated probability spaces. Therefore, we use the short hand notation $(X^x, \tilde{W}^x, \mathbb{Q}^x)$ where X^x is an adapted, continuous stochastic process, $\tilde{W}^x$ is an adapted Brownian motion, both with respect to some filtrated probability space $(\tilde{\Omega}^x, \tilde{\mathcal{F}}^x, \mathbb{Q}^x, (\tilde{\mathcal{F}}_t)^x_{t\geq 0})$, and $(X^x, \tilde{W}^x)$ solves equation (1) with initial value x.

Condition C1. *For each $T > 0$ there exist a measurable $F : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ with*

$$\int_0^T \int_{\mathbb{R}^d} |F(t,x)|^{d+1} \,\mathrm{d}x\mathrm{d}t < \infty$$

and $C_1 = C_1(T), C_2 = C_2(T) \geq 0$ with

$$\int_0^t |B(s,x)|^2 \,\mathrm{d}s \leq \int_0^t |F(s, x(s))| \,\mathrm{d}s + C_1 \sup_{s\in[-r,t]} |x(s)|^2 + C_2$$

for all $t \in [0,T]$and $x \in C\left(\mathbb{R}_{\geq -r}, \mathbb{R}^d\right)$.

Condition C2. *arleta Assume that for all $T > 0$ there exists some $C_\sigma = C_\sigma(T) > 0$ such that*

1. *$C_\sigma^{-1} I_{d\times d} \leq \sigma(t,x)\sigma(t,x)^\top \leq C_\sigma I_{d\times d}$ $\forall t \in [0,T], x \in \mathbb{R}^d$,*
2. *$\|\sigma(t,x) - \sigma(t,y)\|_{HS} \leq C_\sigma |x-y|$ $\forall t \in [0,T], x, y \in \mathbb{R}^d$.*

Condition C3. *Assume that there is an $r_{\tilde{B}} \in (0,r)$ such that*

$$B(t,x) = \tilde{B}(t,x) + b(t, x(t))$$

with $b \in L^{2d+2}\left(\mathbb{R}_{\geq 0} \times \mathbb{R}^d; \mathbb{R}^d\right)$ and $\tilde{B} : \mathbb{R}_{\geq 0} \times C\left(\mathbb{R}_{\geq -r}, \mathbb{R}^d\right) \to \mathbb{R}^d$ measurable where, for fixed $t \geq 0$, $\tilde{B}(t,x)$ depends only on $x_{|[-r, t-r_{\tilde{B}}]}$, i.e.

$$\tilde{B}(t,x) = \tilde{B}(t,y) \text{ if } x(s) = y(s) \ \forall s \in [-r, t - r_{\tilde{B}}].$$

Condition C4. *For $t \in [0,r)$ the function $x \mapsto B(t,x)$ is continuous. Moreover, for each $T>0$ there exist functions $\tilde{F} \in L^{d+1}_{loc}\left([0,T]\times\mathbb{R}^d\right)$ and $G,H:\mathbb{R}_{\geq 0}\to\mathbb{R}_{\geq 0}$ with G monotone increasing and*

$$\lim_{R\to\infty}\frac{H(R)}{R}=\infty$$

such that

$$\int_0^t H\left(|B(s,x)|^2\right)\mathrm{d}s \leq \int_0^t \left|\tilde{F}(s,x(s))\right|\mathrm{d}s + G\left(\sup_{s\in[-r,t]}|x(s)|\right)$$

for all $t\in[0,T]$and $x\in C\left(\mathbb{R}_{\geq -r},\mathbb{R}^d\right)$.

Notation 1.3. In the sequel, let $r>0$ be an arbitrary but fixed number and define

$$\mathcal{C} := C\left([-r,0],\mathbb{R}^d\right)$$

equipped with the supremum norm $\|\cdot\|_\infty$. For a process X defined on $[t-r,t]$ with $t\geq 0$, we write

$$X_t(s) := X(t+s),\ s\in[-r,0].$$

Condition C5. *The non-anticipating function B has bounded memory, i.e. it holds*

$$B(t,x)=B(t,y) \text{ if } x(s)=y(s)\ \forall s\in[t-r,t].$$

Then we use the abuse of notation

$$B(t,x_t)=B(t,x)\ \forall x\in C\left(\mathbb{R}_{\geq -r},\mathbb{R}^d\right)$$

and similarly for $\tilde{B}$ if (C3) *is satisfied.*

The main results read as follows.

Theorem 1.4 (Existence)**.** *Assume* (C1) *and* (C2). *Then for each initial value $x\in\mathcal{C}$, equation* (1) *has a global weak solution $(X^x,\tilde{W}^x,\mathbb{Q}^x)$, which is unique in distribution.*

Theorem 1.5 (Pathwise Uniqueness)**.** *Assume the localized versions of* (C1), (C2) *and* (C3). *Then local pathwise uniqueness holds for equation* (1), *i.e. let (X^x,W) and $(\hat{X}^x,W)$ be two weak solutions of equation* (1) *with initial value $x\in\mathcal{C}$ on some time interval $[0,\tau]$ for some common Brownian motion W and stopping time τ. Then it follows $X^x=\hat{X}^x$ on $[0,\tau]$ almost surely.*

Theorem 1.6 (Strong Feller Property)**.** *Assume* (C1), (C2), (C4) *and* (C5). *Let $(X^x,\tilde{W}^x,\mathbb{Q}^x)$ be weak solutions with initial value $x\in\mathcal{C}$. Then one has the strong Feller property for all $t>r$, i.e.*

$$\lim_{y\to x}\mathbb{E}_{\mathbb{Q}^y}f(X_t^y)=\mathbb{E}_{\mathbb{Q}^x}f(X_t^x)\ \forall f\in B_b(\mathcal{C}).$$

Theorem 1.7 (Stability)**.** *Assume* (C1)*,* (C2)*,*(C3)*,* (C4) *and* (C5)*. Let X^x be the strong solutions with initial value $x \in \mathcal{C}$. Then one has*

$$\lim_{y \to x} \mathbb{E} \, \|X_t^y - X_t^x\|_\infty^\gamma = 0$$

for all $0 < \gamma < 2$ and for $t > r$

$$\lim_{y \to x} \mathbb{E} \, |f(X_t^y) - f(X_t^x)| = 0 \ \forall f \in B_b(\mathcal{C}).$$

Remark 1.8.

1. Conditions (C1) and (C4) are closed under linear combinations.

2. Assume, one has
$$B(t, x_t) = \int_{-r}^{0} k(t, x(t+s)) \mathrm{d}\mu(s)$$
for some Borel measure μ on $[-r, 0]$. Then (C1) is fulfilled if k is of at most linear growth in the second variable uniformly on $[0, r)$ and
$$k \in L^{2d+2}\left([0, T] \times \mathbb{R}^d\right) \ \forall T > 0.$$
If $x \mapsto k(t, x)$ is additionally continuous for $t \in [0, r)$ then condition (C4) will be satisfied. The assumption $\operatorname{supp} \mu \subset [-r, -r_{\tilde{B}}]$ for some $r_{\tilde{B}} \in (0, r)$ implies (C3).

3. The continuity assumption in (C4) is not artificial. Consider the following equation
$$\mathrm{d}X^x(t) = \operatorname{sgn}\left(X^x(t-1)\right) \mathrm{d}t + \mathrm{d}W(t)$$
with
$$\operatorname{sgn}(x) = \begin{cases} 1 & \text{if } x \geq 0, \\ -1 & \text{if } x < 0. \end{cases}$$
Then we have for the strong solutions
$$\begin{aligned} X^0(1) &= W(1) + 1, \\ X^{-1/n}(1) &= W(1) - 1 - \frac{1}{n} \end{aligned}$$
where we denoted constant paths by real numbers. All conditions are fulfilled but the continuity assumption on the interval $[0, r)$. However, neither the strong Feller property nor convergence in probability hold.

4. Consider the one-dimensional, deterministic, functional equation
$$\begin{aligned} \mathrm{d}x(t) &= B(t, x(t-1))\mathrm{d}t, \\ x_0 &= 0 \end{aligned}$$

with

$$B(t,z) := \begin{cases} 8t^7 & \text{if } \ t \in [0,1], z \in \mathbb{R}, \\ \mathbb{1}_{|z|\leq 5} \, |z|^{-1/8} & \text{otherwise.} \end{cases}$$

The drift B fulfills conditions (C1), (C3) and (C4). Computing the solution x yields for $t \in [0,1]$

$$x(t) = t^8$$

Consequently, one has to solve

$$\begin{aligned} \mathrm{d}x(t) &= (t-1)^{-1}\mathrm{d}t \\ x(1) &= 1 \end{aligned}$$

for $t \in [1,2]$. Integrating both sides yields

$$x(t) = 1 + \int_1^t (s-1)^{-1}\mathrm{d}s = \infty$$

for each $t \in (1,2)$. It follows that the equation has no global solution in contrast to its regularized version

$$\mathrm{d}X(t) = B(t, X(t-1))\mathrm{d}t + \mathrm{d}W(t)$$

although the conditions (C1), (C3) and (C4) are fulfilled.

2. A-priori Estimates and Existence

In the sequel, denote by M^x, $x \in \mathcal{C}$ the global, unique strong solution of

$$\begin{aligned} \mathrm{d}M^x(t) &= \sigma\left(t, M^x(t)\right)\mathrm{d}W(t), \\ M_0^x &= x. \end{aligned}$$

Notation 2.1. We denote by $\|\cdot\|_{op}$ and $\|\cdot\|_{HS}$ the operator norm and respectively theHilbert-Schmidt norm for matrices $A \in \mathbb{R}^{d\times d}$, i.e.

$$\|A\|_{op} = \sup_{v\in\mathbb{R}^d, |v|=1} |Av| \, , \ \|A\|_{HS} = \sqrt{\sum_{i,j=1}^{d} |A^{i,j}|^2}.$$

Additionally, we write for $a, b \in [-\infty, +\infty]$

$$a \wedge b := \min\{a,b\}, \ \ a \vee b := \max\{a,b\}.$$

Remark 2.2. Condition (C2) implies the following inequalities

$$\|\sigma\|_{op} \vee \left\|\sigma^{-1}\right\|_{op} \leq \sqrt{C_\sigma}.$$

Lemma 2.3. *Assume* (C2) *and let* $T > 0$, $p > \frac{d+2}{2}$ *be given. Then one has for all* $0 \leq S < T$ *and* $f \in L_p\left([S,T] \times \mathbb{R}^d\right)$ *the estimate*

$$\mathbb{E}\left(\int_S^T f(t, M^x(t))\mathrm{d}t \middle| \mathcal{F}_S\right) \leq C \|f\|_{L^p\left([S,T]\times\mathbb{R}^d\right)}$$

for some constant $C = C(d, p, T, C_\sigma)$. *In particular, the constant* C *is independent of the initial value* $x \in \mathcal{C}$.

Proof. This follows directly from Theorem 2.1 in [17]. □

Lemma 2.4. *Assume* (C2). *Then for any* $R, T > 0$ *and* $p > \frac{d+2}{2}$ *there exists a constant* $C_R = C_R(d, p, T, C_\sigma)$ *such that*

$$\mathbb{E}\exp\left(\int_0^T f(t, M^x(t))\mathrm{d}t\right) \leq C_R$$

for all $f \in L^p\left([0,T] \times \mathbb{R}^d\right)$ *with* $\|f\|_{L^p\left([0,T]\times\mathbb{R}^d\right)} \leq R$.

Proof. See Lemma 2.1 in [17]. □

Lemma 2.5. *Assume* (C2). *Then for any* $T > 0$ *and* $0 \leq \alpha < (2dC_\sigma T)^{-1}$, *it holds*

$$\mathbb{E}\exp\left(\alpha \sup_{0\leq t\leq T} |M^x(t)|^2\right) \leq \frac{4}{\sqrt{1-2\alpha d C_\sigma T}} \exp\left(\frac{\alpha}{1-2\alpha d C_\sigma T} |x(0)|^2\right).$$

Proof. See Lemma 2.4 in [1]. □

Let X be a d-dimensional Itō-process of the form

$$\mathrm{d}X(t) = b(t)\mathrm{d}t + \sigma(t)\mathrm{d}W(t)$$

where W is a d-dimensional Brownian motion. Set

$$a^{ij}(t) := \frac{1}{2}\sigma(t)\sigma(t)^\top$$

and let τ_R be the first exit time of $X(t)$ from the ball B_R.

Lemma 2.6 (Krylov's Estimate). *For every stopping time* γ *and nonnegative Borel function* $f : \mathbb{R}_{\geq 0} \times \mathbb{R}^d \to \mathbb{R}_{\geq 0}$ *one has*

$$\begin{aligned}&\mathbb{E}\int_0^{\gamma\wedge\tau_R} (\det a(t))^{\frac{1}{d+1}} f(t, X(t))\mathrm{d}t \\ \leq & N(d)\left(\mathbb{B}^2 + \mathbb{A}\right)^{\frac{d}{2(d+1)}} \left(\int_0^\infty \int_{|x|\leq R} f^{d+1}(t,x)\mathrm{d}x\mathrm{d}t\right)^{\frac{1}{d+1}}\end{aligned}$$

where

$$\mathbb{A} := \mathbb{E}\int_0^{\gamma\wedge\tau_R} \operatorname{tr} a(t)\mathrm{d}t, \quad \mathbb{B} := \mathbb{E}\int_0^{\gamma\wedge\tau_R} |b(t)|\,\mathrm{d}t$$

and $N(d)$ *is a constant depending only on the dimension* d.

Corollary 2.7. *Assume* (C1), (C2) *and let* $T > 0$. *Furthermore, let* $(X^x, \tilde{W}^x, \mathbb{Q}^x)$ *be a solution of equation* (1) *on some time interval* $[-r, \tau]$ *where* τ *is some stopping time with* $0 \leq \tau \leq T$. *Then one has*

$$\mathbb{E} \int_0^\tau |B(t, X^x)|^2 \, \mathrm{d}t \leq 2C_1 \mathbb{E}\left[\sup_{-r \leq t \leq \tau} |X^x(t)|^2\right] + C$$

where $C = C\left(d, T, C_2, C_\sigma, \|F\|_{L^{d+1}([0,T]\times\mathbb{R}^d)}\right)$ *is some constant.*

Proof. The proof is similar to the one of Corollary 3.2. in [7]. By Krylov's estimate and Young's inequality, one has

$$\begin{aligned}
&\mathbb{E} \int_0^\tau |B(t, X^x)|^2 \, \mathrm{d}t \\
\leq &\mathbb{E} \int_0^\tau |F(t, X^x(t))| \, \mathrm{d}t + C_1 \mathbb{E}\left[\sup_{-r \leq t \leq \tau} |X^x(t)|^2\right] + C_2 \\
\leq &N(d) \left(\mathbb{B}^2 + \mathbb{A}\right)^{\frac{d}{2(d+1)}} \|F\|_{L^{d+1}([0,T]\times\mathbb{R}^d)} + C_1 \mathbb{E}\left[\sup_{-r \leq t \leq \tau} |X^x(t)|^2\right] + C_2 \\
\leq &\frac{1}{2T} \left(\mathbb{B}^2 + \mathbb{A}\right) + C_1 \mathbb{E}\left[\sup_{-r \leq t \leq \tau} |X^x(t)|^2\right] + \tilde{C} \\
\leq &\frac{1}{2} \mathbb{E} \int_0^\tau |B(t, X^x)|^2 \, \mathrm{d}t + C_1 \mathbb{E}\left[\sup_{-r \leq t \leq \tau} |X^x(t)|^2\right] + C
\end{aligned}$$

where C and $\tilde{C}$ are constants depending only on d, T, C_2, C_σ and $\|F\|_{L^{d+1}([0,T]\times\mathbb{R}^d)}$ with

$$\mathbb{A} := \frac{1}{2} \mathbb{E} \int_0^\tau \operatorname{tr}\left(\sigma(t, X^x(t))\sigma(t, X^x(t))^\top\right) \mathrm{d}t, \quad \mathbb{B} := \mathbb{E} \int_0^\tau |B(t, X^x)| \, \mathrm{d}t.$$

□

Corollary 2.8. *Assume* (C1), (C2) *and let* $T > 0$. *Furthermore, let* $(X^x, \tilde{W}^x, \mathbb{Q}^x)$ *be a solution of equation* (1) *on some time interval* $[-r, \tau]$ *where* τ *is some stopping time with* $0 \leq \tau \leq T$. *Then one has*

$$\mathbb{E}\left[\sup_{-r \leq t \leq \tau} |X^x(t)|^2\right] \leq C\left(1 + \|x\|_\infty^2\right)$$

where $C = C\left(d, T, C_1, C_2, C_\sigma, \|F\|_{L^{d+1}([0,T]\times\mathbb{R}^d)}\right)$ *is some constant.*

Proof. Applying Gronwall's lemma and Doob's maximal inequality. □

Corollary 2.9. *Assume* (C1), (C2) *and let* $T > 0$. *Moreover, let* $(X^x, \tilde{W}^x, \mathbb{Q}^x)$ *be a weak solution of equation* (1) *on* $[-r, \tau]$ *for some stopping time* $0 \leq \tau \leq T$. *Then for any Borel function* $f : \mathbb{R}^{d+1} \to \mathbb{R}_{\geq 0}$ *and* $q \geq d+1$, *one has*

$$\mathbb{E} \int_0^{T\wedge\tau} f(t, X^x(t)) \mathrm{d}t \leq N \|f\|_{L^q(T)}$$

where $N = N\left(d, T, C_1, C_2, C_\sigma, \|F\|_{L^{d+1}([0,T]\times\mathbb{R}^d)}, \|x\|_\infty\right)$ *is a constant.*

Proof. This follows directly from Krylov's estimate and the Corollaries before. □

Theorem 2.10. *Assume* (C1) *and* (C2). *Then for every initial values $x \in \mathcal{C}$, equation* (1) *has a global weak solution. Moreover, for each weak solution $(X^x, \tilde{W}^x, \mathbb{Q}^x)$ of equation* (1) *on some time interval $[-r, T]$, $T > 0$, one has*

$$\mathbb{Q}^x_{X^x}(A) = \mathbb{E}_{\mathbb{P}}\left[\mathbb{1}_A(M^x)\exp\left(\int_0^T a^x(t)^\top \mathrm{d}W(t) - \frac{1}{2}\int_0^T |a^x(t)|^2\,\mathrm{d}t\right)\right],$$
$$a^x(t) := \sigma(t, M^x(t))^{-1}B(t, M^x),\ t \in [0,T]$$

for all measurable $A \subset C([-r,T], \mathbb{R}^d)$.

Proof. At first, we show the existence of a weak solution. The strong solution M^x is by definition $(\mathcal{F}_t)_{t\geq 0}$-adapted where $(\mathcal{F}_t)_{t\geq 0}$ is the augmented filtration generated by W. Next, we construct a probability measure on

$$\mathcal{F}_\infty := \sigma\left(\mathcal{F}_t : t \geq 0\right)$$

such that M^x is a global weak solution for equation (1). By Lemma 2.4, Lemma 2.5, conditions (C1) and (C2), there exist for each $T > 0$ a partition $0 = T_0 \leq T_1, \cdots \leq T_{n-1} \leq T_n = T$, $n \in \mathbb{N}$ with

$$\mathbb{E}_{\mathbb{P}}\exp\left(\frac{1}{2}\int_{T_{i-1}}^{T_i}\left|\sigma(t, M^x(t))^{-1}B(t, M^x)\right|^2\mathrm{d}t\right) < \infty,\ i = 1, \ldots, n.$$

Therefore, Novikov's condition is fulfilled for each subinterval, which gives that

$$\begin{aligned} t \mapsto \exp\Bigg(&\int_{t\wedge T_{i-1}}^{t\wedge T_i}\left(\sigma(s, M^x(s))^{-1}B(s, M^x)\right)^\top \mathrm{d}W(s) \\ &- \frac{1}{2}\int_{t\wedge T_{i-1}}^{t\wedge T_i}|\sigma(s, M^x(s))B(s, M^x)|^2\,\mathrm{d}s\Bigg)\end{aligned}$$

is a martingale for $i = 1, \ldots, n$. Consequently,

$$\begin{aligned} t \mapsto \exp\Bigg(&\int_0^t\left(\sigma(s, M^x(s))^{-1}B(s, M^x)\right)^\top \mathrm{d}W(s) \\ &- \frac{1}{2}\int_0^t|\sigma(t, M^x(s))B(t, M^x)|^2\,\mathrm{d}s\Bigg)\end{aligned}$$

is a martingale and by Girsanov's theorem,

$$\bar{W}(t) := W(t) - \int_0^t \sigma(s, M^x(s))^{-1}B(s, M^x)\mathrm{d}s,\ t \geq 0$$

is a Brownian motion on $[0, T]$ under the probability measure

$$\begin{aligned} \mathrm{d}\bar{\mathbb{P}}_T := \exp\Bigg(&\int_0^T\left(\sigma(t, M^x(t))^{-1}B(t, M^x)\right)^\top \mathrm{d}W(t) \\ &- \frac{1}{2}\int_0^T|\sigma(t, M^x(t))B(t, M^x)|^2\,\mathrm{d}t\Bigg)\mathrm{d}\mathbb{P}\end{aligned}$$

and $(M^x, \bar{W}, \bar{\mathbb{P}}_T)$ is a weak solution of (1) on $[-r, T]$ for each $T > 0$. Additionally, one has for $0 < T_1 < T_2$

$$\bar{\mathbb{P}}_{T_1}(A) = \bar{\mathbb{P}}_{T_2}(A) \ \forall A \in \mathcal{F}_{T_1},$$

so the probability measure on $\mathcal{F}_\infty$ uniquely defined by

$$\bar{\mathbb{P}}(A) := \mathbb{P}_T(A) \ \forall T > 0, A \in \mathcal{F}_T$$

is indeed well-defined and $(M^x, \bar{W}, \bar{\mathbb{P}})$ is a global weak solution.

Now, let $(X^x, \tilde{W}^x, \mathbb{Q}^x)$ be a weak solution on some time interval $[0, T]$, $T > 0$. The following approach is inspired by the techniques used in [11]. Define

$$\tau^n(\omega) := \inf\left\{t \geq 0 : \int_0^t |B(s, \omega)|^2 \, \mathrm{d}s \geq n\right\} \wedge T, \ \omega \in C([-r, T], \mathbb{R}^d), \ n \in \mathbb{N}.$$

Then the stopped process $X^{x,n}(t) := X^x(t \wedge \tau^n(X^x))$, $t \in [-r, T]$ fulfills the equation

$$\mathrm{d}X^{x,n}(t) = \mathbb{1}_{t \leq \tau^n(X^{x,n})} B(t, X^{x,n})\mathrm{d}t + \mathbb{1}_{t \leq \tau^n(X^{x,n})} \sigma(t, X^{x,n}(t))\mathrm{d}\tilde{W}^x$$

By construction, Novikov's condition is fulfilled. Consequently, Girsanov's theorem is applicable and

$$\tilde{W}^{x,n}(t) := \int_0^{t \wedge \tau^n(X^{x,n})} \sigma(s, X^{x,n}(s))^{-1} B(s, X^{x,n})\mathrm{d}s + \tilde{W}^x(t), \ t \geq 0$$

is a Brownian motion with respect to the probability measure

$$\begin{aligned}\mathrm{d}\mathbb{Q}^{x,n} := \exp\Bigg(&- \int_0^{\tau^n(X^{x,n})} \left(\sigma(t, X^{x,n}(t))^{-1} B(t, X^{x,n})\right)^\top \mathrm{d}\tilde{W}^x(t) \\ &- \frac{1}{2} \int_0^{\tau^n(X^{x,n})} |\sigma(t, X^{x,n}(t)) B(t, X^{x,n})|^2 \, \mathrm{d}t \Bigg) \mathrm{d}\mathbb{Q}^x.\end{aligned}$$

The process $X^{x,n}$ solves the equation

$$\begin{aligned}\mathrm{d}X^{x,n}(t) &= \sigma(t, X^{x,n}(t))\mathrm{d}\tilde{W}^{x,n}(t), \ t \in [0, \tau^n(X^{x,n})], \\ X_0^{x,n} &= x.\end{aligned}$$

Such a solution is (locally) pathwise unique , i.e.

$$X^{x,n}(t) = M^{x,n}(t), \ t \in [-r, \tau^n(X^{x,n})]$$

where $M^{x,n}$ is the unique strong solution of

$$\begin{aligned}\mathrm{d}M^{x,n}(t) &= \sigma(t, M^{x,n}(t))\mathrm{d}\tilde{W}^{x,n}(t), \\ M_0^{x,n} &= x.\end{aligned}$$

and it holds

$$\tau^n(X^{x,n}) = \tau^n(M^{x,n}) \text{ a.s.}$$

Moreover, $\mathbb{Q}^x$ and $\mathbb{Q}^{x,n}$ are equivalent. Thus,

$$\begin{aligned}
&\mathbb{Q}^x(X^x \in A)\\
&= \lim_{n\to\infty} \mathbb{Q}^x(\tau^n(X^x) = T, X^x \in A)\\
&= \lim_{n\to\infty} \mathbb{Q}^x(\tau^n(X^{x,n}) = T, X^{x,n} \in A)\\
&= \lim_{n\to\infty} \mathbb{E}_{\mathbb{Q}^{x,n}}\left[\mathbb{1}_{\tau^n(X^{x,n})=T}\mathbb{1}_A(X^{x,n}) \exp\left(\int_0^T \left(\sigma(t, X^{x,n}(t))^{-1}B(t, X^{x,n})\right)^\top \mathrm{d}\tilde{W}^{x,n}(t)\right.\right.\\
&\quad \left.\left. - \frac{1}{2}\int_0^T \left|\sigma(t, X^{x,n}(t))^{-1}B(t, X^{x,n})\right|^2 \mathrm{d}t\right)\right]\\
&= \lim_{n\to\infty} \mathbb{E}_{\mathbb{Q}^{x,n}}\left[\mathbb{1}_{\tau^n(M^{x,n})=T}\mathbb{1}_A(M^{x,n}) \exp\left(\int_0^T \left(\sigma(t, M^{x,n}(t))^{-1}B(t, M^{x,n})\right)^\top \mathrm{d}\tilde{W}^{x,n}(t)\right.\right.\\
&\quad \left.\left. - \frac{1}{2}\int_0^T \left|\sigma(t, M^{x,n}(t))^{-1}B(t, M^{x,n})\right|^2 \mathrm{d}t\right)\right]\\
&= \lim_{n\to\infty} \mathbb{E}_{\mathbb{P}}\left[\mathbb{1}_{\tau^n(M^x)=T}\mathbb{1}_A(M^x) \exp\left(\int_0^T \left(\sigma(t, M^x(t))^{-1}B(t, M^x)\right)^\top \mathrm{d}W^x(t)\right.\right.\\
&\quad \left.\left. - \frac{1}{2}\int_0^T \left|\sigma(t, M^x(t))^{-1}B(t, M^x)\right|^2 \mathrm{d}t\right)\right]\\
&=\mathbb{E}_{\mathbb{P}}\left[1_A(M^x) \exp\left(\int_0^T \left(\sigma(t, M^x(t))^{-1}B(t, M^x)\right)^\top \mathrm{d}W^x(t)\right.\right.\\
&\quad \left.\left. - \frac{1}{2}\int_0^T \left|\sigma(t, M^x(t))^{-1}B(t, M^x)\right|^2 \mathrm{d}t\right)\right]
\end{aligned}$$

for all measurable $A \subset C([-r, T], \mathbb{R}^d)$. $\square$

Lemma 2.11. *Assume* (C1) *with* $C_1 = 0$, (C2) *and let* $T > 0$, $q \geq d+1$ *be given. Moreover, let* $(X^x, \tilde{W}^x, \mathbb{Q}^x)$ *be a weak solution of equation* (1) *on* $[-r, \tau]$ *for some stopping time* $0 \leq \tau \leq T$. *Then one has*

$$\sup_{f\in L^q([0,T]\times\mathbb{R}^d):\|f\|_{L^q}\leq R} \mathbb{E}_{\mathbb{Q}^x} \exp\left(\int_0^\tau f(t, X^x(t))\mathrm{d}t\right) < \infty$$

for all $R > 0$.

Proof. Let

$$a^x(t) := \sigma(t, M^x(t))^{-1}B(t, M^x),\ t \in [0, T].$$

Analogous proceeding as in proof of Theorem 2.10 gives

$$
\begin{aligned}
&\mathbb{E}_{\mathbb{Q}^x} \exp\left(\int_0^\tau f(t, X^x(t))\mathrm{d}t\right) \\
\leq &\mathbb{E}_{\mathbb{P}} \exp\left(\int_0^T f(t, M^x(t))\mathrm{d}t + \int_0^T a^x(t)^\top \mathrm{d}W(t) - \frac{1}{2}\int_0^T |a^x(t)|^2 \,\mathrm{d}t\right) \\
\leq &\left[\mathbb{E}_{\mathbb{P}} \exp\left(\int_0^T 2f(t, M^x(t))\mathrm{d}t\right)\right]^{\frac{1}{2}} \left[\mathbb{E}_{\mathbb{P}} \exp\left(2\int_0^T a^x(t)^\top \mathrm{d}W(t) - \int_0^T |a^x(t)|^2 \,\mathrm{d}t\right)\right]^{\frac{1}{2}} \\
\leq &\left[\mathbb{E}_{\mathbb{P}} \exp\left(\int_0^T 2f(t, M^x(t))\mathrm{d}t\right)\right]^{\frac{1}{2}} \cdot \left[\mathbb{E}_{\mathbb{P}} \exp\left(6\int_0^T |a^x(t)|^2 \,\mathrm{d}t\right)\right]^{\frac{1}{4}}.
\end{aligned}
$$

The uniform bound follows from condition (C1) and Lemma 2.4. □

Lemma 2.12. *Assume* (C1) *with* $C_1 = 0$, (C2) *and and let* $T > 0$ *be given. Let* $(X^x, \tilde{W}^x, \mathbb{Q}^x)$ *be a weak solution of equation* (1) *on* $[-r, \tau]$ *for some stopping time* $0 \leq \tau \leq T$. *Then the following inequality holds.*

$$
\begin{aligned}
&\mathbb{E}_{\mathbb{Q}^x} \exp\left(\alpha \sup_{-r\leq t\leq \tau} |X^x(t)|^2\right) \\
\leq &\frac{C}{\sqrt[4]{1 - 4\alpha d C_\sigma T}} \exp\left(\frac{\alpha}{1 - 4\alpha d C_\sigma T} \|x\|_\infty^2\right)
\end{aligned}
$$

for all $0 \leq \alpha < (4dC_\sigma T)^{-1}$ *and a constant* $C = C\left(d, T, C_2, C_\sigma, \|F\|_{L^{d+1}([0,T]\times\mathbb{R}^d)}\right)$.

Proof. As before, let

$$
a^x(t) := \sigma(t, M^x(t))^{-1} B(t, M^x),\ t \in [0, T].
$$

By the assumed conditions and Lemma 2.4, one has

$$
\begin{aligned}
&\mathbb{E}_{\mathbb{P}} \exp\left(6\int_0^T |a^x(t)|^2 \,\mathrm{d}t\right) \\
\leq &K_1 \mathbb{E}_{\mathbb{P}} \exp\left(6C_\sigma \int_0^T |B(t, M^x)|^2 \,\mathrm{d}t\right) \\
\leq &K_2 \mathbb{E}_{\mathbb{P}} \exp\left(6C_\sigma \int_0^T |F(t, M^x(t))| \,\mathrm{d}t\right) \\
\leq &K_3
\end{aligned}
$$

for constants K_1, K_2 and K_3 that only depend on d, T, C_2, C_σ, $\|F\|_{L^{d+1}([0,T]\times\mathbb{R}^d)}$. By

Theorem 2.10, one obtains

$$
\begin{aligned}
&\mathbb{E}_{\mathbb{Q}^x}\left(\alpha \sup_{-r\le t\le \tau} |X^x(t)|^2\right)\\
\le&\mathbb{E}_{\mathbb{P}} \exp\left(\alpha \sup_{-r\le t\le T} |M^x(t)|^2 + \int_0^T a^x(t)^\top \mathrm{d}W(t) - \frac{1}{2}\int_0^T |a^x(t)|^2 \,\mathrm{d}t\right)\\
\le&\left[\mathbb{E}_{\mathbb{P}} \exp\left(2\alpha \sup_{-r\le t\le T} |M^x(t)|^2\right)\right]^{\frac{1}{2}} \cdot \left[\mathbb{E}_{\mathbb{P}} \exp\left(2\int_0^T a^x(t)^\top \mathrm{d}W(t) - \int_0^T |a^x(t)|^2 \,\mathrm{d}t\right)\right]^{\frac{1}{2}}\\
\le&\left[\mathbb{E}_{\mathbb{P}} \exp\left(2\alpha \sup_{-r\le t\le T} |M^x(t)|^2\right)\right]^{\frac{1}{2}} \cdot \left[\mathbb{E}_{\mathbb{P}} \exp\left(6\int_0^T |a^x(t)|^2 \,\mathrm{d}t\right)\right]^{\frac{1}{4}}\\
\le&\frac{C}{\sqrt[4]{1-4\alpha d C_\sigma T}} \exp\left(\frac{\alpha}{1-4\alpha d C_\sigma T}\|x\|_\infty^2\right)
\end{aligned}
$$

for a constant $C = C\left(d, T, C_2, C_\sigma, \|F\|_{L^{d+1}([0,T]\times\mathbb{R}^d)}\right)$. □

3. Strong Feller Property

The following theorem is a consequence of a log-Harnack inequality that has been shown in [16] and requires the Lipschitz-continuity of σ in space.

Theorem 3.1. *Assume* (C2)*. Then one has for all* $t > r$

$$\lim_{y\to x} \mathbb{E} f(M_t^y) = \mathbb{E} f(M_t^x)\ \forall f \in B_b(\mathcal{C}).$$

Lemma 3.2. *Assume* (C1), (C2), (C4) *and* (C5)*. Then one has*

$$\lim_{y\to x} \mathbb{P}\left(\int_0^T |B(t, M_t^x) - B(t, M_t^y)|^2 \,\mathrm{d}t > \varepsilon\right) = 0\ \forall \varepsilon > 0.$$

Proof. By Theorems 3.1 and 1.1, one has for all $t > r$

$$\lim_{y\to x} \mathbb{E}\,|f(M_t^x) - f(M_t^y)| = 0\ \forall f \in B_b(\mathcal{C}).$$

Therefore, one has for all $f \in B_b([0,T]\times\mathcal{C})$

$$\lim_{y\to x} \mathbb{E}\int_r^T |f(t, M_t^x) - f(t, M_t^y)| \,\mathrm{d}t = 0.$$

Consequently,

$$\lim_{y\to x} \mathbb{P}\otimes\lambda_{|[r,T]}\left(|B(\cdot, M_\cdot^y) - B(\cdot, M_\cdot^x)| > \varepsilon\right) = 0\ \forall \varepsilon > 0.$$

By condition (C4), one has also

$$\lim_{y\to x} \mathbb{P}\otimes\lambda_{|[0,r]}\left(|B(\cdot,M^y_\cdot)-B(\cdot,M^x_\cdot)|>\varepsilon\right)=0\ \forall\varepsilon>0.$$

Therefore, it holds

$$\lim_{y\to x} \mathbb{P}\otimes\lambda_{|[0,T]}\left(|B(\cdot,M^y_\cdot)-B(\cdot,M^x_\cdot)|>\varepsilon\right)=0\ \forall\varepsilon>0.$$

Now, define for $R>0$

$$B^R(t,x):=\mathbb{1}_{\sup\limits_{-r\le s\le t}|x(s)|^2<R}B(t,x_t),\ x\in C\left(\mathbb{R}_{\ge -r},\mathbb{R}^d\right)$$

Each B^R, $R>0$ fulfills condition (C1) and the suitably modified version of (C4) with bounded G. It follows

$$\sup_{y\in\mathcal{C}}\mathbb{E}\int_0^T H\left(\left|B^R(t,M^y)\right|^2\right)\mathrm{d}t<\infty,\ R>0$$

by Lemma 2.3, 2.5 and condition (C4). Hence, $\left\{\left|B^R(\cdot,M^y_\cdot)\right|^2 : y\in\mathcal{C}\right\}$ is uniformly integrable for each $R>0$. Since

$$\lim_{y\to x}\mathbb{E}\left\|M^y_t-M^x_t\right\|_\infty\ \forall t\ge 0,$$

it holds

$$\lim_{y\to x} \mathbb{P}\otimes\lambda_{|[0,T]}\left(\left|B^R(\cdot,M^y)-B^R(\cdot,M^x)\right|>\varepsilon\right)=0\ \forall\varepsilon>0.$$

Furthermore,

$$\lim_{R\to\infty}\sup_{x\in\mathcal{C},\|x\|\le\tilde{R}}\mathbb{P}\left(\sup_{t\in[-r,T]}|M^x(t)|\ge R\right)=0\ \forall\tilde{R}>0$$

holds by condition (C2). Thus,

$$\begin{aligned}
&\limsup_{y\to x}\mathbb{P}\left(\int_0^T|B(t,M^x(t))-B(t,M^y(t))|^2\,\mathrm{d}t>\varepsilon\right)\\
\le&\limsup_{y\to x}\mathbb{P}\left(\int_0^T\left|B^R(t,M^x(t))-B^R(t,M^y(t))\right|^2\mathrm{d}t>\varepsilon\right)\\
&+2\sup_{z\in\mathcal{C},\|z\|\le 2\|z\|_\infty}\mathbb{P}\left(\sup_{t\in[-r,T]}|M^z(t)|\ge R\right)\\
=&2\sup_{z\in\mathcal{C},\|z\|\le 2\|z\|_\infty}\mathbb{P}\left(\sup_{t\in[-r,T]}|M^z(t)|\ge R\right)
\end{aligned}$$

Now, one can let $R\to\infty$, which proofs the claim. □

Proof of Theorem 1.6. Let $t > r$ and $f \in B_b(\mathcal{C})$, then one has by Theorem 2.10

$$\begin{aligned}
&\mathbb{E}_{\mathbb{Q}^x} f(X_t^x) - \mathbb{E}_{\mathbb{Q}^y} f(X_t^y) \\
=&\mathbb{E}_{\mathbb{P}}(D^x(t) f(M_t^x)) - \mathbb{E}_{\mathbb{P}}(D^y(t) f(M_t^y)) \\
=&\mathbb{E}_{\mathbb{P}}[D^x(t)(f(M_t^x) - f(M_t^y))] + \mathbb{E}_{\mathbb{P}}[(D^x(t) - D^y(t)) f(M_t^y)] \\
\leq&\mathbb{E}_{\mathbb{P}}[D^x(t)(f(M_t^x) - f(M_t^y))] + \|f\|_\infty \mathbb{E}_{\mathbb{P}} |D^x(t) - D^y(t)|
\end{aligned}$$

where we define for every $z \in \mathcal{C}$

$$\begin{aligned}
a^z(t) &:= \sigma(t, M^z(t))^{-1} B(t, M_t^z), \\
D^z(t) &:= \exp\left(\int_0^t a^z(s)^\top \mathrm{d}W(s) - \frac{1}{2} \int_0^t |a^z(s)|^2 \,\mathrm{d}s \right).
\end{aligned}$$

By condition (C2), Itō's formula and the stochastic Gronwall Lemma A.5, it holds

$$\lim_{y \to x} \mathbb{P}\left(|M_t^y - M_t^x| > \varepsilon \right) = 0 \ \forall \varepsilon > 0. \tag{2}$$

Applying Theorems 1.1 and 3.1, gives

$$\lim_{y \to x} \mathbb{E}_{\mathbb{P}} |f(M_t^y) - f(M_t^x)| = 0$$

and in particular,

$$\lim_{y \to x} \mathbb{P}\left(|D^x(t) f(M_t^{x_n}) - D^x(t) f(M_t^x)| > \varepsilon \right) = 0 \ \forall \varepsilon > 0.$$

By the dominated convergence theorem, it follows

$$\lim_{y \to x} \mathbb{E}_{\mathbb{P}}[D^x(t)(f(M_t^y) - f(M_t^x))] = 0.$$

Consequently, it remains to show that

$$\lim_{y \to x} \mathbb{E}_{\mathbb{P}} |D^y(t) - D^x(t)| = 0.$$

Since one has $\mathbb{E}_{\mathbb{P}} D^z(t) = 1$ for all $z \in \mathcal{C}$, it suffices to show

$$\lim_{y \to x} \mathbb{P}\left(|D^y(t) - D^x(t)| > \varepsilon \right) = 0 \ \forall \varepsilon > 0$$

by standard measure theoretic arguments. Therefore, it is sufficient to show

$$\lim_{y \to x} \mathbb{P}\left(\int_0^t |a^y(s) - a^x(s)|^2 \,\mathrm{d}s > \varepsilon \right) = 0 \ \forall \varepsilon > 0$$

by the martingale isometry. One has

$$\begin{aligned}
&\int_0^t |a^y(s) - a^x(s)|^2 \,\mathrm{d}s \\
\leq& 2 \int_0^t \left\| \sigma\left(s, M^y(s)\right)^{-1} - \sigma\left(s, M^x(s)\right)^{-1} \right\|_{op}^2 |B(s, M_s^x)|^2 \,\mathrm{d}s \\
&+ 2 C_\sigma \int_0^t |B(s, M_s^y) - B(s, M_s^x)|^2 \,\mathrm{d}t.
\end{aligned}$$

The second term converges to zero by the assumed conditions and Lemma 3.2. Moreover,

$$\lim_{y\to x} \mathbb{P}\otimes\lambda_{|[0,t]}\left(\left\|\sigma\left(\cdot, M^y(\cdot)\right)^{-1} - \sigma\left(\cdot, M^x(\cdot)\right)^{-1}\right\|_{op} > \varepsilon\right) = 0\ \forall\varepsilon > 0$$

holds by (2), the continuity of σ in space and the continuity of the inverting map $A \mapsto A^{-1}$ on the space of invertible matrices. Additionally, one can bound the first integrand by

$$2C_\sigma \left|B(\cdot, M_\cdot^x)\right|^2,$$

which is $\mathbb{P}\otimes\lambda_{|[0,t]}$-integrable by Lemma 2.3. Consequently, one can apply the dominated convergence theorem and the proof is complete. □

4. Pathwise Uniqueness and Stability

Notation 4.1. We introduce - as in [17]- the following function space. For $p, \in (1,\infty)$ and $0 \le S < T$, denote by $W_p^{1,2}\left([S,T]\times\mathbb{R}^d\right)$ the closure of compactly supported, smooth functions on $[S,T]\times\mathbb{R}^d$ with respect to the norm

$$\|u\|_{W_p^{1,2}\left([S,T]\times\mathbb{R}^d\right)} := \|\partial_t u\|_{L^p[S,T]} + \|u\|_{L^p([S,T];W^{2,p})},\ u\in C_c^\infty([S,T]\times\mathbb{R}^d).$$

Let $p := 2d+2$. By Theorem A.1, for every $0 < T \le T_0$, there exists a solution

$$\tilde{u}(\cdot;T) \in \left(W_p^{1,2}\left([0,T_0]\times\mathbb{R}^d\right)\right)^d$$

of the coordinatewise PDE system

$$\begin{aligned} \partial_t\tilde{u}(t,x;T) + L_t\tilde{u}(t,x;T) + b(t,x) = 0,\\ \tilde{u}(T,x;T) = 0 \end{aligned}$$

for all $t\in[0,T]$ and $x\in\mathbb{R}^d$ where

$$L_t v(t,x) := \frac{1}{2}\sum_{i,j,k=1}^d \sigma^{i,k}(t,x)\sigma^{j,k}(t,x)\partial_i\partial_j v(t,x) + b(t,x)\cdot\nabla v(t,x),\ v\in W_p^{1,2}\left([0,T_0]\times\mathbb{R}^d\right).$$

Additionally, it holds

$$\sup_{T\in[0,T_0]} \left\|\tilde{u}^i(\cdot;T)\right\|_{W_p^{1,2}\left([0,T]\times\mathbb{R}^d\right)} < \infty,\ i = 1,\dots,d$$

and by the embedding Theorem A.2, there exists a uniform δ such that for all $0\le S\le T$ with $T - S\le\delta$

$$|\tilde{u}(t,x;T) - \tilde{u}(t,y;T)| \le \frac{1}{2}|x-y|$$

for all $t\in[S,T]$ and $x,y\in\mathbb{R}^d$. Furthermore, the function

$$u(t,x;T) := \tilde{u}(t,x;T) + x$$

satisfies coordinatewise the equation

$$\begin{aligned}\partial_t u(t,x;T) + L_t u(t,x;T) &= 0,\\ u(T,x;T) &= x.\end{aligned}$$

Proof of Theorem 1.5. Let (X^x, W) and $(\hat{X}^x, W)$ be two weak solutions of equation (1) with initial value $x \in \mathcal{C}$ for some common Brownian motion W on the time interval $[0,\tau]$ for some stopping time τ. By localization, we can assume that condition (C1) is fulfilled with $C_1 = 0$ and that τ is bounded by some $T_0 > 0$. Choose $\delta > 0$ like above with the additional restraint $\delta < r_{\tilde{B}}$. By induction, it suffices to prove for every $0 \leq S \leq T \leq T_0$ with $T - S \leq \delta$

$$\begin{aligned}X^x_{|[-r,S\wedge\tau]} &= \hat{X}^x_{|[-r,S\wedge\tau]}\\ \Longrightarrow X^x_{T\wedge\tau} &= \hat{X}^x_{T\wedge\tau}.\end{aligned}$$

For the sake of simplicity, we write $u(\cdot) := u(\cdot;T)$. Furthermore, define

$$\begin{aligned}Y(t) &:= u(t, X(t)),\ S\wedge\tau \leq t \leq T\wedge\tau,\\ \hat{Y}(t) &:= u(t, \hat{X}^x(t)),\ S\wedge\tau \leq t \leq T\wedge\tau.\end{aligned}$$

By the choice of δ, one has for the difference processes $Z(t) := X^x(t) - \hat{X}^x(t)$ and $\tilde{Z}(t) := Y(t) - \hat{Y}(t)$

$$\frac{1}{2}\left|\tilde{Z}(t)\right| \leq |Z(t)| \leq \frac{3}{2}\left|\tilde{Z}(t)\right|,\ S\wedge\tau \leq t \leq T\wedge\tau.$$

Due to Lemma 2.11, Lemma A.3 is applicable, which gives for $S\wedge\tau \leq t \leq T\wedge\tau$

$$\begin{aligned}\tilde{Z}(t) =& \int_S^t \left(Du(s,X^x(s))\tilde{B}(s,X^x) - Du(s,\hat{X}^x(s))\tilde{B}(s,\hat{X}^x)\right)\mathrm{d}s\\ &+ \int_S^t \left(Du(s,X^x(s))\sigma(s,X^x(s)) - Du(s,\hat{X}^x(s))\sigma(s,\hat{X}^x(s))\right)\mathrm{d}W(s)\\ =& \int_S^t \left(Du(s,X^x(s))\tilde{B}(s,X^x) - Du(s,\hat{X}^x(s))\tilde{B}(s,X^x)\right)\mathrm{d}s\\ &+ \int_S^t \left(Du(s,X^x(s))\sigma(s,X^x(s)) - Du(s,\hat{X}^x(s))\sigma(s,\hat{X}^x(s))\right)\mathrm{d}W(s)\end{aligned}$$

and consequently

$$\begin{aligned}&\mathrm{d}\left|\tilde{Z}\right|^2(t)\\ =&2\tilde{Z}(t)^\top\left(Du(t,X^x(t))\tilde{B}(t,X^x) - Du(t,\hat{X}^x(t))\tilde{B}(t,X^x)\right)\mathrm{d}t\\ &+ 2\tilde{Z}(t)^\top\left(Du(t,X^x(t))\sigma(t,X^x(t)) - Du(t,X^{x_n}(t))\sigma(t,X^{x_n}(t))\right)\mathrm{d}W(t)\\ &+ \left\|Du(t,X^x(t))\sigma(t,X^x(t)) - Du(t,\hat{X}^x(t))\sigma(t,\hat{X}^x(t))\right\|^2_{HS}\mathrm{d}t\end{aligned}$$

Using the boundedness of Du, condition (C1) and Young's inequality gives for $S \le t_1 \le t_2 \le T$

$$
\begin{aligned}
&\left|\tilde{Z}(t_2)\right|^2 - \left|\tilde{Z}(t_1)\right|^2 \\
\le &\int_{t_1}^{t_2} \left|\tilde{Z}(s)\right| \|Du(s, X^x(s)) - Du(s, X^{x_n}(s))\|_{op} \left|\tilde{B}(s, X^x)\right| \mathrm{d}s \\
+c &\int_{t_1}^{t_2} \tilde{Z}(s)^\top \left(Du(s, X^x(s))\sigma(s, X^x(s)) - Du(s, X^{x_n}(s))\sigma(s, X^{x_n}(s))\right) \mathrm{d}W(s) \\
+c &\int_{t_1}^{t_2} \|Du(s, X^x(s))\sigma(s, X^x(s)) - Du(s, X^{x_n}(s))\sigma(s, X^{x_n}(s))\|_{HS}^2 \,\mathrm{d}s
\end{aligned}
$$

where $c > 0$ is a constant. As in [6], one can use a suitable multiplier of the form $\exp(-A(t))$ where A is an adapted, continuous process. Here, we choose

$$
\begin{aligned}
A(t) := c &\int_S^t \left|\tilde{B}(s, X^x)\right| \frac{\left\|Du(s, X^x(s)) - Du(s, \hat{X}^x(s))\right\|_{op}}{\left|\tilde{Z}(s)\right|} \mathbb{1}_{\tilde{Z}(s)\neq 0} \mathrm{d}s \\
+ c &\int_S^t \frac{\left\|Du(s, X^x(s))\sigma(s, X^x(s)) - Du(s, \hat{X}^x(s))\sigma(s, \hat{X}^x(s))\right\|_{HS}^2}{\left|\tilde{Z}(s)\right|^2} \mathbb{1}_{\tilde{Z}(s)\neq 0} \mathrm{d}s
\end{aligned}
$$

for $S \le t \le T$. To show that A is indeed well defined - namely finite - it suffices to show

$$
\mathbb{E} \exp\left(\frac{1}{2} A(T)\right) < \infty.
$$

Since u belongs coordinatewise to $W_p^{1,2}\left([0, T_0] \times \mathbb{R}^d\right)$ and by conditions (C2), it holds

$$
(Du \cdot \sigma)^{i,j} \in L^p\left(T_0; W^{1,p}\left(\mathbb{R}^d\right)\right), \ i, j = 1, \dots, d.
$$

Additionally, $C_c^\infty\left(\mathbb{R}^{d+1}\right)$ is dense in $L^p\left(T_0; W^{1,p}\left(\mathbb{R}^d\right)\right)$. Hence, by Young's inequality, Lemma 2.11 and Lemma 2.12, it suffices to show for all $\tilde{R} > 0$ the existence of a constant $C_R > 0$ such that

$$
\mathbb{E} \exp\left(\int_S^T \frac{\left|f(s, X^x(s)) - f(s, \hat{X}^x(s))\right|^2}{\left|\tilde{Z}(s)\right|^2} \mathbb{1}_{\tilde{Z}(s)\neq 0} \mathrm{d}s\right) \le C_{\tilde{R}}
$$

for all $f \in C^\infty\left(\mathbb{R}^{d+1}\right)$ with $\|f\|_{L^p\left(T_0; W^{1,p}\left(\mathbb{R}^d\right)\right)} \le R$. By Lemmas 2.11 and A.4, one obtains

$$\begin{aligned}
&\mathbb{E}\exp\left(\int_S^T \frac{\left|f(s,X^x(s)) - f(s,\hat{X}^x(s))\right|^2}{\left|\tilde{Z}(s)\right|^2} \mathbb{1}_{\tilde{Z}(s)\neq 0}\mathrm{d}s\right)\\
\leq &\mathbb{E}\exp\left(C_d^2\int_S^T \left(\mathcal{M}\left|\nabla f\right|(X^x(s)) + \mathcal{M}\left|\nabla f\right|(\hat{X}^x(s))\right)^2 \mathrm{d}s\right)\\
<&\infty.
\end{aligned}$$

Now, it holds for $S \leq t \leq T$

$$e^{-A(t)}\left|\tilde{Z}(t)\right|^2 \leq c\int_S^t e^{-A(s)}\left|\tilde{Z}(s)\right|^2 \mathrm{d}s + \text{local martingale}$$

by the Itō formula. Applying the stochastic Gronwall Lemma A.5 gives

$$\mathbb{E}\left[\sup_{t\in[S,T]} e^{-\frac{1}{2}A(t)}\left|\tilde{Z}(t)\right|\right] = 0,$$

which finishes the proof. □

The following result is a rather technical one, which will be used to proof Theorem 1.7.

Proposition 4.2. *Assume* (C1), (C2), (C3), (C4) *and* (C5). *Furthermore, let* X^x, $x\in\mathcal{C}$ *be the strong solutions to equation* (1) *with initial value* x *and assume that*

$$\lim_{y\to x}\mathbb{P}\left(\left\|X_S^y - X_S^x\right\|_\infty > \varepsilon\right) = 0 \;\forall\varepsilon>0, \forall x\in\mathcal{C} \tag{3}$$

for some $S\geq 0$. *Then one has for each* $R>0$

$$\lim_{n\to\infty}\mathbb{E}\int_{S\wedge\tau_R^{x,y}}^{(S+r_{\tilde{B}})\wedge\tau_R^{x,y}} \left|\tilde{B}(s,X_s^y) - \tilde{B}(s,X_s^x)\right|^2 \mathrm{d}s = 0$$

where

$$\tau_R^{x,y} := \sup\left\{t\geq 0 : \sup_{-r\leq s\leq t}\left|X^x(s)\right|^2 < R,\ \sup_{-r\leq s\leq t}\left|X^y(s)\right|^2 < R\right\}.$$

Proof. By condition (C3), one can write

$$\tilde{B}(t,X_t^x) = g(t,X_S^x)\ t\in[S,S+r_{\tilde{B}}], x\in\mathcal{C}. \tag{4}$$

If $S>r$, Theorem 1.6 gives

$$\lim_{y\to x}\mathbb{E}f(X_S^y) = \mathbb{E}f(X_S^x)\ \forall f\in B_b(\mathcal{C}).$$

Consequently, combining it with (3), (4) and Theorem 1.1 gives

$$\lim_{y\to x} \mathbb{P}\left(\left|g(t, X_S^y) - g(t, X_S^x)\right| > \varepsilon\right) = 0 \;\forall \varepsilon > 0, t \in [S, S + r_{\tilde{B}}]. \tag{5}$$

If $S \le r$, one can use the continuity assumption (C4) and (3) to deduce (5), too. Therefore

$$\lim_{y\to x} \mathbb{P}\otimes\lambda_{|[S,S+r_{\tilde{B}}]}\left(|B(\cdot, X_\cdot^y) - B(\cdot, X_\cdot^x)|^2 > \varepsilon\right) = 0 \;\forall \varepsilon > 0.$$

It follows

$$\lim_{y\to x} \mathbb{P}\otimes\lambda_{|[S,S+r_{\tilde{B}}]}\left(\mathbb{1}_{\tau_R^{x,y}\le\cdot}\, |B(\cdot, X_\cdot^y) - B(\cdot, X_\cdot^x)|^2 > \varepsilon\right) = 0 \;\forall \varepsilon > 0.$$

Now, one can use Lemma 2.3 and condition (C4) to obtain

$$\sup_{y\in\mathcal{C}} \mathbb{E}\int_{S\wedge\tau_R^{x,y}}^{(S+r_{\tilde{B}})\wedge\tau_R^{x,y}} H\left(|B(t, X_t^y)|^2\right) \mathrm{d}t < \infty,$$

which guarantees the uniform integrability of $\left\{\mathbb{1}_{\tau_R^{y,y}<\cdot}\, |B(\cdot, M_\cdot^y)|^2 : y\in\mathcal{C}\right\}$ with respect to the measure $\mathbb{P}\otimes\lambda_{|[S,S+r_{\tilde{B}}]}$. □

Proof of Theorem 1.7. Choose $\delta > 0$ like before with the additional restraint $\delta < r_{\tilde{B}}$. By induction and Lemma (2.12), it suffices to prove for every $0 \le S \le T \le T_0$ with $T - S \le \delta$ the implication

$$\begin{aligned}
&\lim_{y\to x} \mathbb{E}\left\|X_S^y - X_S^x\right\|_\infty^\gamma = 0 \;\forall x\in\mathcal{C}, 0<\gamma<2\\
\implies &\lim_{y\to x} \mathbb{P}\left(\sup_{s\in[S,T]} |X^y(s) - X^x(s)| > \varepsilon\right) = 0 \;\forall \varepsilon > 0, x\in\mathcal{C}.
\end{aligned}$$

For the sake of simplicity, we write $u(\cdot) := u(\cdot; T)$. Furthermore, define

$$\begin{aligned}
Y^x(t) &:= u(t, X(t)),\; S\le t\le T,\\
Y^y(t) &:= u(t, X^y(t)),\; S\le t\le T.
\end{aligned}$$

By the choice of δ, one has for the difference processes $Z(t) := X^x(t) - X^y(t)$ and $\tilde{Z}(t) := Y^x(t) - Y^y(t)$

$$\frac{1}{2}\left|\tilde{Z}(t)\right| \le |Z(t)| \le \frac{3}{2}\left|\tilde{Z}(t)\right|,\; S\le t\le T.$$

Due to Lemma 2.11, Lemma A.3 is applicable, which gives

$$\begin{aligned}
\tilde{Z}(t) = &\int_S^t \left(Du(s, X^x(s))\tilde{B}(s, X_s^x) - Du(s, X^y(s))\tilde{B}(s, X_s^y)\right)\mathrm{d}s\\
&+ \int_S^t \left(Du(s, X^x(s))\sigma(s, X^x(s)) - Du(s, X^y(s))\sigma(s, X^y(s))\right)\mathrm{d}W(s)
\end{aligned}$$

and consequently

$$\begin{aligned}
&\mathrm{d}\left|\tilde{Z}\right|^2(t)\\
=&2\tilde{Z}(t)^\top\left(Du(t,X^x(t))\tilde{B}(t,X^x_t)-Du(t,X^y(t))\tilde{B}(t,X^y_t)\right)\mathrm{d}t\\
&+2\tilde{Z}(t)^\top\left(Du(t,X^x(t))\sigma(t,X^x(t))-Du(t,X^y(t))\sigma(t,X^y(t))\right)\mathrm{d}W(t)\\
&+\|Du(t,X^x(t))\sigma(t,X^x(t))-Du(t,X^y(t))\sigma(t,X^y(t))\|_{HS}^2\,\mathrm{d}t
\end{aligned}$$

Using the boundedness of Du and Young's inequality gives for $S\le t_1\le t_2\le T$

$$\begin{aligned}
&\left|\tilde{Z}(t_2)\right|^2-\left|\tilde{Z}(t_1)\right|^2\\
\le&c\int_{t_1}^{t_2}\left|\tilde{Z}(s)\right|^2\mathrm{d}s\\
&+c\int_{t_1}^{t_2}\left|\tilde{B}(s,X^x_s)-\tilde{B}(s,X^y_s)\right|^2\mathrm{d}s\\
&+c\int_{t_1}^{t_2}\left|\tilde{Z}(s)\right|\|Du(s,X^x(s))-Du(s,X^y(s))\|_{op}\left|\tilde{B}(s,X^x_s)\right|\mathrm{d}s\\
&+c\int_{t_1}^{t_2}\tilde{Z}(s)^\top\left(Du(s,X^x(s))\sigma(s,X^x(s))-Du(s,X^y(s))\sigma(s,X^y(s))\right)\mathrm{d}W(s)\\
&+c\int_{t_1}^{t_2}\|Du(s,X^x(s))\sigma(s,X^x(s))-Du(s,X^y(s))\sigma(s,X^y(s))\|_{HS}^2\,\mathrm{d}s
\end{aligned}$$

where $c>0$ is a constant. Like before, we one can use the multiplier $\exp(-A(t))$ where

$$\begin{aligned}
A(t):=&\,c\int_S^t|B(s,X^x_s)|\frac{\|Du(s,X^x(s))-Du(s,X^y(s))\|_{op}}{\left|\tilde{Z}(s)\right|}\mathbb{1}_{\tilde{Z}(s)\neq0}\mathrm{d}s\\
&+c\int_S^t\frac{\|Du(s,X^x(s))\sigma(s,X^x(s))-Du(s,X^y(s))\sigma(s,X^y(s))\|_{HS}^2}{\left|\tilde{Z}(s)\right|^2}\mathbb{1}_{\tilde{Z}(s)\neq0}\mathrm{d}s
\end{aligned}$$

for $S\le t\le T$. Again, one has

$$\mathbb{E}\exp\left(\frac{1}{2}A(T)\right)\le\hat{C}$$

where $\hat{C}$ is some constant not depending on $x,y\in\mathcal{C}$. By the Itō formula, it holds for $S\le t\le T$

$$\begin{aligned}
e^{-A(t)}\left|\tilde{Z}(t)\right|^2\le&\left|\tilde{Z}(S)\right|^2+c\int_S^t\left|\tilde{B}(s,X^x_s)-\tilde{B}(s,X^y_s)\right|^2\mathrm{d}s\\
&+c\int_S^t e^{-A(s)}\left|\tilde{Z}(s)\right|^2\mathrm{d}s+\text{local martingale.}
\end{aligned}$$

Applying the stochastic Gronwall Lemma A.5 gives

$$\mathbb{E}\left[\sup_{t\in[S,T]} e^{-\frac{1}{2}A(t)}\left|\tilde{Z}(t)\right|\right] \le \tilde{C}\mathbb{E}\left|\tilde{Z}(S)\right| + \tilde{C}\mathbb{E}\left(\int_S^T \left|\tilde{B}(s,X_s^x)-\tilde{B}(s,X_s^y)\right|^2 \mathrm{d}s\right)^{\frac{1}{2}}$$

for a constant $\tilde{C}$ which does not depend on $x, y \in \mathcal{C}$. By Lemma 2.8,

$$\lim_{R\to\infty} \sup_{z\in\mathcal{C},\|z\|_\infty\le 2\|x\|_\infty} \mathbb{P}\left(\sup_{-r\le t\le T} |X^x(t)|^2 > R\right) = 0.$$

holds. Thus, applying Lemma 4.2 and the induction hypothesis gives

$$\lim_{y\to x} \mathbb{P}\left(\int_S^T \left|\tilde{B}(s,X_s^x)-\tilde{B}(s,X_s^y)\right|^2 \mathrm{d}s > \varepsilon\right) = 0 \ \forall \varepsilon > 0.$$

By Corollary 2.8, one has

$$\sup_{z\in\mathcal{C},\|z\|_\infty\le 2\|x\|_\infty} \mathbb{E}\int_S^T \left|\tilde{B}(s,X_s^x)-\tilde{B}(s,X_s^y)\right|^2 \mathrm{d}s < \infty$$

and consequently,

$$\lim_{y\to x} \mathbb{E}\left(\int_S^T \left|\tilde{B}(s,X_s^x)-\tilde{B}(s,X_s^y)\right|^2 \mathrm{d}s\right)^{\frac{1}{2}} = 0.$$

□

A. Appendix

Theorem A.1. *Assume* (C2) *and* $b \in L^p\left([0,T]\times\mathbb{R}^d\right)$ *with* $p > d+2$. *Then for any* $T > 0$ *and* $f \in L^p\left([0,T]\times\mathbb{R}^d\right)$, *there exists a unique solution* $u \in W_p^{1,2}\left([0,T]\times\mathbb{R}^d\right)$ *of the following PDE*

$$\begin{aligned} \partial_t u(t,x) + L_t u(t,x) + f(t,x) &= 0, \\ u(T,x) &= 0 \end{aligned}$$

with the bound

$$\|u\|_{W_p^{1,2}\left([S,T]\times\mathbb{R}^d\right)} \le C\|f\|_{L^p\left([S,T]\times\mathbb{R}^d\right)}$$

for any $S \in [0,T]$ *and some constant* $C = C(T, C_\sigma, p, \|b\|_{L^p\left([0,T]\times\mathbb{R}^d\right)}) > 0$.

Proof. See Theorem 10.3 in [10]. □

Theorem A.2. *Let* $p \in (1,\infty)$, $T > 0$ *and* $u \in W_p^{1,2}\left([0,T]\times\mathbb{R}^d\right)$.

1. *If $p > \frac{d+2}{2}$, then u is a bounded Hölder continuous function on $[0,T] \times \mathbb{R}^d$ and for any $0 < \varepsilon,\ \delta \leq 1$ satisfying*

$$\varepsilon + \frac{d+2}{p} < 2, \quad 2\delta + \frac{d+2}{p} < 2,$$

   there exists a constant $N = N(p, \varepsilon, \delta)$ such that

$$|u(t,x) - u(s,x)| \leq N\,|t-s|^{\delta}\, \|u\|^{1-\frac{1}{q}-\delta}_{L^p\left(T;\mathrm{W}^{2,p}\left(\mathbb{R}^d\right)\right)} \|\partial_t u\|^{\frac{1}{p}+\delta}_{L^p\left([0,T]\times\mathbb{R}^d\right)},$$
$$|u(t,x)| + \frac{|u(t,x)-u(t,y)|}{|x-y|^{\varepsilon}} \leq NT^{-\frac{1}{p}}\left(\|u\|_{L^p\left(T;\mathrm{W}^{2,p}\left(\mathbb{R}^d\right)\right)} + T\,\|\partial_t u\|_{L^p\left([0,T]\times\mathbb{R}^d\right)}\right)$$

   for all $s,t \in [0,T]$ and $x,y \in \mathbb{R}^d, x \neq y$.

2. *If $p > d+2$, then ∇u is a bounded Hölder continuous function on $[0,T] \times \mathbb{R}^d$ and for any $\varepsilon \in (0,1)$ satisfying*

$$\varepsilon + \frac{d+2}{p} < 1,$$

   there exists a constant $N = N(p,\varepsilon)$ such that

$$|\nabla u(t,x) - \nabla u(s,x)| \leq N\,|t-s|^{\delta}\, \|u\|^{1-\frac{1}{p}-\frac{\varepsilon}{2}}_{L^p\left(T;\mathrm{W}^{2,p}\left(\mathbb{R}^d\right)\right)} \|\partial_t u\|^{\frac{1}{p}+\frac{\varepsilon}{2}}_{L^p\left([0,T]\times\mathbb{R}^d\right)},$$
$$|\nabla u(t,x)| + \frac{|\nabla u(t,x)-\nabla u(t,y)|}{|x-y|^{\varepsilon}} \leq NT^{-\frac{1}{p}}\left(\|u\|_{L^p\left(T;\mathrm{W}^{2,p}\left(\mathbb{R}^d\right)\right)} + T\,\|\partial_t u\|_{L^p\left([0,T]\times\mathbb{R}^d\right)}\right)$$

   for all $s,t \in [0,T]$ and $x,y \in \mathbb{R}^d, x \neq y$.

Proof. See [5, p. 22, 23, 36]. □

In the next lemma we identify every $u \in W^{1,2}_p$ with its regular version.

Lemma A.3 (Itō formula for $W^{1,2}_p$-functions)**.** *Let $T > 0$, $p > d+2$. Let $X : \Omega\times[0,T] \to \mathbb{R}^d$ be a semimartingale on some filtrated probability space $\left(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0}\right)$ of the form*

$$\mathrm{d}X(t) = b(t)\mathrm{d}t + \sigma(t)\mathrm{d}W(t)$$

where W is a d-dimensional Brownian motion, $b : \Omega \times [0,T] \to \mathbb{R}^d$ and $\sigma : \Omega \times [0,T] \to \mathbb{R}^{d\times d}$ are progressively measurable with

$$\mathbb{P}\left(\|b\|_{L^1[0,T]} + \left\|a^{i,j}\right\|_{L^{\delta}[0,T]} < \infty\right) = 1,\ i,j = 1,\ldots,d$$

for some $1 < \delta \leq \infty$ where $a := \sigma\sigma^{\top}$. Furthermore, assume that there exists a constant $C > 0$ with

$$\mathbb{E}\int_0^T f(t,X(t))\mathrm{d}t \leq C\,\|f\|_{L^{p/\delta *}\left([0,T]\times\mathbb{R}^d\right)}$$

for all $f \in L^{p/\delta^*}\left([0,T]\times\mathbb{R}^d\right)$ *where* δ^* *denotes the conjugate exponent of* δ*. Then for any* $u \in W_p^{1,2}\left([0,T]\times\mathbb{R}^d\right)$*, the Itō formula holds, i.e.*

$$\begin{aligned} u(t,X(t)) - u(0,X(0)) &= \int_0^t \partial_t u(s,X(s))\mathrm{d}s + \int_0^t \nabla u(s,X(s))^\top b(s)\mathrm{d}s \\ &\quad + \int_0^t \nabla u(s,X(s))^\top \sigma(s)\mathrm{d}W(s) \\ &\quad + \frac{1}{2}\sum_{i,j=1}^d \int_0^t \partial_i\partial_j u(s,X(s))a^{i,j}(s)\mathrm{d}s. \end{aligned}$$

Proof. See [1]. □

Let ϕ be a locally integrable function on $\mathbb{R}^d$. The Hardy-Littlewood maximal function is defined by

$$\mathcal{M}\phi(x) := \sup_{0<r<\infty} \frac{1}{|B_r|}\int_{B_r} \phi(x+y)\mathrm{d}y$$

where B_r is the Euclidean ball of radius r. The following result is cited from Appendix A in [3].

Lemma A.4.

1. *There exists a constant* $C_d > 0$ *such that for all* $\phi \in C^\infty\left(\mathbb{R}^d\right)$ *and* $x, y \in \mathbb{R}^d$*,*

$$|\phi(x) - \phi(y)| \leq C_d\,|x-y|\left(\mathcal{M}\,|\nabla\phi|\,(x) + \mathcal{M}\,|\nabla\phi|\,(y)\right).$$

2. *For any* $p > 1$*, there exists a constant* $C_{d,p}$ *such that for all* $\phi \in L^p\left(\mathbb{R}^d\right)$*,*

$$\|\mathcal{M}\phi\|_{L^p} \leq C_{d,p}\,\|\phi\|_{L^p.}$$

For a real-valued process denote $Y^*(t) := \sup_{0\leq s\leq t} Y(s)$.

Lemma A.5. *Let* Z *and* H *be nonnegative, adapted processes with continuous paths and assume that* ψ *is nonnegative and progressively measurable. Let* M *be a continuous local martingale starting at 0. If*

$$Z(t) \leq \int_0^t \psi(s)Z(s)\mathrm{d}s + M(t) + H(t)$$

holds for all $t \geq 0$*, then for* $p \in (0,1)$ *and* $\mu, \nu > 1$ *such that* $\frac{1}{\mu} + \frac{1}{\nu} = 1$ *and* $p\nu < 1$*, we have*

$$\mathbb{E}\sup_{0\leq s\leq t} Z(s)^p \leq (c_{p\nu}+1)^{1/\nu}\left(\mathbb{E}\exp\left\{p\mu\int_0^t \psi(s)\mathrm{d}s\right\}\right)^{1/\mu}\left(\mathbb{E}(H^*(t))^{p\nu}\right)^{1/\nu}$$

where

$$c_p := \left(4 \wedge \frac{1}{p}\right)\frac{\pi p}{\sin(\pi p)}.$$

If ψ is deterministic, then

$$\mathbb{E} \sup_{0 \le s \le t} Z(s)^p \le (1 + c_p) \exp\left\{ p \int_0^t \psi(s) \mathrm{d}s \right\} (\mathbb{E}(H^*(t))^p)$$

and

$$\mathbb{E}Z(t) \le \exp\left\{ \int_0^t \psi(s) \mathrm{d}s \right\} \mathbb{E}H^*(t).$$

Proof. See [14]. □

References

[1] S. Bachmann. Well-posedness and stability for a class of stochastic delay differential equations with singular drift. *Stochastics and Dynamics*, 0(0):1850019, 0.

[2] S. Bachmann. On the Strong Feller Property of Stochastic Delay Differential Equations with Singular Drift. *ArXiv e-prints*, September 2017.

[3] Gianluca Crippa and Camillo De Lellis. Estimates and regularity results for the DiPerna-Lions flow. *J. Reine Angew. Math.*, 616:15–46, 2008.

[4] A. Es-Sarhir, M.-K. von Renesse, and M. Scheutzow. Harnack inequality for functional sdes with bounded memory. *Electron. Commun. Probab.*, 14:560–565, 2009.

[5] E. Fedrizzi. Uniqueness and flow theorems for solutions of SDEs with low regularity of the drift. *tesi di Laurea in Matematica, Università di Pisa*, 2009.

[6] E. Fedrizzi and F. Flandoli. Pathwise uniqueness and continuous dependence of SDEs with non-regular drift. *Stochastics*, 83(3):241–257, 2011.

[7] I. Gyöngy and T. Martínez. On stochastic differential equations with locally unbounded drift. *Czechoslovak Math. J.*, 51(126)(4):763–783, 2001.

[8] X. Huang. Strong solutions for functional sdes with singular drift. *Stochastics and Dynamics*, 0(0):1850015, 0.

[9] N. V. Krylov. Estimates of the maximum of the solution of a parabolic equation and estimates of the distribution of a semimartingale. *Mat. Sb. (N.S.)*, 130(172)(2):207–221, 284, 1986.

[10] N. V. Krylov and M. Röckner. Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Related Fields*, 131(2):154–196, 2005.

[11] R. S. Liptser and A. N. Shiryaev. *Statistics of Random Processes: I. General Theory*, pages 286–297. Springer, 2001.

[12] B. Maslowski and J. Seidler. Probabilistic approach to the strong feller property. *Probability Theory and Related Fields*, 118(2):187–210, Oct 2000.

[13] N. I. Portenko. *Generalized diffusion processes*, volume 83 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1990. Translated from the Russian by H. H. McFaden.

[14] M. Scheutzow. A stochastic gronwall lemma. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 16(02):1350019, 2013.

[15] A. Ju. Veretennikov. Strong solutions of stochastic differential equations. *Teor. Veroyatnost. i Primenen.*, 24(2):348–360, 1979.

[16] F.-Y. Wang and C. Yuan. Harnack inequalities for functional SDEs with multiplicative noise and applications. *Stochastic Process. Appl.*, 121(11):2692–2710, 2011.

[17] X. Zhang. Stochastic homeomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients. *Electron. J. Probab.*, 16:no. 38, 1096–1116, 2011.

[18] A. K. Zvonkin. A transformation of the phase space of a diffusion process that will remove the drift. *Mat. Sb. (N.S.)*, 93(135):129–149, 152, 1974.

Bibliography

[1] S. Bachmann. Well-posedness and stability for a class of stochastic delay differential equations with singular drift. *Stochastics and Dynamics*, 0(0):1850019, 0.

[2] S. Bachmann. On the Strong Feller Property of Stochastic Delay Differential Equations with Singular Drift. *ArXiv e-prints*, September 2017.

[3] S. Bachmann. On the Strong Feller Property and Well-Posedness for SDEs with Functional, Locally Unbounded Drift. *ArXiv e-prints*, August 2018.

[4] R. Catellier and M. Gubinelli. Averaging along irregular curves and regularisation of odes. *Stochastic Processes and their Applications*, 126, 02 2016.

[5] J. Cortissoz. On the Skorokhod Representation Theorem. *ArXiv Mathematics e-prints*, January 2006.

[6] G. Crippa and C. De Lellis. Estimates and regularity results for the DiPerna-Lions flow. *J. Reine Angew. Math.*, 616:15–46, 2008.

[7] A. M. Davie. Uniqueness of solutions of stochastic differential equations. *International Mathematics Research Notices*, 2007, 10 2007.

[8] A. M. Davie. *Individual Path Uniqueness of Solutions of Stochastic Differential Equations*, pages 213–225. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.

[9] R. M. Dudley. *Real Analysis and Probability*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2 edition, 2002.

[10] O. Enchev and D. W. Stroock. Rademacher's theorem for Wiener functionals. *Ann. Probab.*, 21(1):25–33, 1993.

[11] A. Es-Sarhir, M.-K. von Renesse, and M. Scheutzow. Harnack inequality for functional sdes with bounded memory. *Electron. Commun. Probab.*, 14:560–565, 2009.

[12] E. Fedrizzi. Uniqueness and flow theorems for solutions of SDEs with low regularity of the drift. *tesi di Laurea in Matematica, Università di Pisa*, 2009.

[13] E. Fedrizzi and F. Flandoli. Pathwise uniqueness and continuous dependence of SDEs with non-regular drift. *Stochastics*, 83(3):241–257, 2011.

[14] F. Flandoli, M. Gubinelli, and E. Priola. Flow of diffeomorphisms for SDEs with unbounded Hölder continuous drift. *Bull. Sci. Math.*, 134(4):405–422, 2010.

[15] B Gess and M. Maurelli. Well-posedness by noise for scalar conservation laws. *arXiv e-prints*, page arXiv:1701.05393, January 2017.

[16] I. Gyöngy and T. Martínez. On stochastic differential equations with locally unbounded drift. *Czechoslovak Math. J.*, 51(126)(4):763–783, 2001.

[17] X. Huang. Strong solutions for functional sdes with singular drift. *Stochastics and Dynamics*, 0(0):1850015, 0.

[18] N. V. Krylov. Estimates of the maximum of the solution of a parabolic equation and estimates of the distribution of a semimartingale. *Mat. Sb. (N.S.)*, 130(172)(2):207–221, 284, 1986.

[19] N. V. Krylov and M. Röckner. Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Related Fields*, 131(2):154–196, 2005.

[20] R. S. Liptser and A. N. Shiryaev. *Statistics of Random Processes: I. General Theory*, pages 286–297. Springer, 2001.

[21] B. Maslowski and J. Seidler. Probabilistic approach to the strong feller property. *Probability Theory and Related Fields*, 118(2):187–210, Oct 2000.

[22] T. Meyer-Brandis and F. Proske. Construction of strong solutions of sde's via malliavin calculus. *Journal of Functional Analysis*, 258(11):3922 – 3953, 2010.

[23] D. Nualart and Y. Ouknine. Regularization of differential equations by fractional noise. *Stochastic Processes and their Applications*, 102(1):103 – 116, 2002.

[24] A. Pilipenko and F. N. Proske. On a selection problem for small noise perturbation in the multidimensional case. *Stochastics and Dynamics*, 18(06):1850045, 2018.

[25] N. I. Portenko. *Generalized diffusion processes*, volume 83 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1990. Translated from the Russian by H. H. McFaden.

[26] Ivan L. Reilly. Quasi-gauge spaces. *Journal of the London Mathematical Society*, s2-6(3):481–487, 1973.

[27] M. Scheutzow. A stochastic gronwall lemma. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 16(02):1350019, 2013.

[28] D. Shakhmatov and J. Spěvák. Group-valued continuous functions with the topology of pointwise convergence. *Topology and its Applications*, 157(8):1518 – 1540, 2010. Advances in Set-Theoretic Topology: Proceedings of the Conference in Honour of Professor Tsugunori Nogura on his Sixtieth Birthday (9–19 June 2008, Erice, Sicily, Italy).

[29] J.A. van Casteren. *Markov Processes, Feller Semigroups and Evolution Equations*. Series on concrete and applicable mathematics. World Scientific, 2011.

[30] A. W. van der Vaart and J. Wellner. *Weak Convergence and Empirical Processes: With Applications to Statistics.* Springer Series in Statistics. Springer, 1996.

[31] A. Ju. Veretennikov. Strong solutions of stochastic differential equations. *Teor. Veroyatnost. i Primenen.*, 24(2):348–360, 1979.

[32] M.-K. von Renesse and M. Scheutzow. Existence and uniqueness of solutions of stochastic functional differential equations. *Random Oper. Stoch. Equ.*, 18(3):267–284, 2010.

[33] F.-Y. Wang and C. Yuan. Harnack inequalities for functional SDEs with multiplicative noise and applications. *Stochastic Process. Appl.*, 121(11):2692–2710, 2011.

[34] S. Willard. *General Topology.* Addison-Wesley series in mathematics. Dover Publications, 1970.

[35] T. Yamada and S. Watanabe. On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.*, 11(1):155–167, 1971.

[36] X. Zhang. Strong solutions of sdes with singular drift and sobolev diffusion coefficients. *Stochastic Processes and their Applications*, 115(11):1805 – 1818, 2005.

[37] X. Zhang. Stochastic homeomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients. *Electron. J. Probab.*, 16:no. 38, 1096–1116, 2011.

[38] X. Zhang. Stochastic differential equations with Sobolev drifts and driven by α-stable processes. *Ann. Inst. Henri Poincaré Probab. Stat.*, 49(4):1057–1079, 2013.

[39] A. K. Zvonkin. A transformation of the phase space of a diffusion process that will remove the drift. *Mat. Sb. (N.S.)*, 93(135):129–149, 152, 1974.

Selbständigkeitserklärung

Hiermit erkläre ich, die vorliegende Dissertation selbständig und ohne unzulässige fremde Hilfe angefertigt zu haben. Ich habe keine anderen als die angeführten Quellen und Hilfsmittel benutzt und sämtliche Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder unveröffentlichten Schriften entnommen wurden, und alle Angaben, die auf mündlichen Auskünften beruhen, als solche kenntlich gemacht. Ebenfalls sind alle von anderen Personen bereitgestellten Materialen oder erbrachten Dienstleistungen als solche gekennzeichnet.

Leipzig, 11. Dezember 2019

Stefan Bachmann

Wissenschaftlicher Werdegang

2008-2015	Mathematikstudium an der Universität Leipzig mit Nebenfach Informatik und Abschluss Dipl.-Math.
2015-2019	Wissenschaftlicher Mitarbeiter an der Mathematischen Fakultät Leipzig